Cover picture

Halley's Comet as represented in the
Bayeux Tapestry An article on
modern comet theory appears on
page 15

John R. Freeman

Volume 263   September 2 1976

Cover picture

"Whale fishing". Part of a facsimile
of a woodcut in the *Cosmographie
Universelle* of Thevet (Paris, 1574).
See page 91.
Mary Evans Picture Library

Volume 263   September 9 1976

Guide to authors

● Articles are up to 3,000 words in length with at most six displayed items (figures and tables) and may either be reports of major research developments in a subject or broader reviews of progress.

● Letters are brief reports of research of unusual and wide interest, not in general longer than 1,000 words; at most they have three or four displayed items (figures and tables).

● 'Matters Arising' permits occasional short discussion of papers that have previously appeared in *Nature*. There is a limit of 300 words.

Manuscripts may be submitted either to London or Washington. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should conform to the Système International. Greek characters should be identified in the margin on their first appearance. Equations should occupy single lines if possible, exp (a) is preferred to e" if 'a' is more than one character. Articles should be accompanied by an abstract of not more than fifty words listing the main conclusions.

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) To refer to several references by the same author at once, only one reference number need be given.

(ii) Cite first and last pages.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth ed. (Butterworth, 1963–65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated. 'Personal communication' and 'unpublished work' should be incorporated in the text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings to be in Indian ink on heavy cartridge paper, tracing paper or similar materials: thin, shiny, torn or folded material should be avoided. Scales for graphs should be chosen such that simple figures will not occupy an inordinate amount of space. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text they should contain only essential material. Ideally, an unlabelled original and a labelled copy should be provided: labelling on photographs should, if possible, be avoided entirely. Records, charts etc. which will not reproduce well should be redrawn or simply described in the text. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

A fuller guide appeared in *Nature* (246, 238; 1973).

Cover picture

Chris Sherwell reports from the new
National Air and Space Museum on
page 180.

Volume 263   September 16 1976

MERCK SHARP AND DOHME RESEARCH LABORATORIES
SCIENTIFIC LECTURE 1976
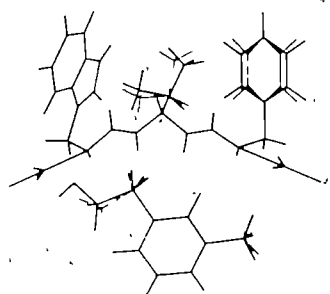
FLEXIBLE DRUG MOLECULES AND DYNAMIC RECEPTORS

by PROFESSOR R. J. P. WILLIAMS, F.R.S.

Wadham College and the Inorganic Chemistry Laboratory, Oxford

The lecture will be presented at the Royal College of Physicians, London, on Wednesday, 20th October 1976, at 5 p.m. (Tea and biscuits will be served from 4 p.m.). Invitations are now being sent and may be obtained by completing the application form.

To: Dr. W. G. Thomas,
    Merck Sharp & Dohme Research Laboratories,
    Hertford Road, Hoddesdon, Herts., EN11 9BU

For the 1976 lecture, please send an invitation to:

Name

Address

Cover picture

Ardnamurchan, a Tertiary igneous
ring complex. Evolution of a ring
complex is discussed on page 306.

Guide to authors

● Articles are up to 3,000 words in length with at most six displayed items (figures and tables) and may either be reports of major research developments in a subject or broader reviews of progress.

● Letters are brief reports of research of unusual and wide interest, not in general longer than 1,000 words; at most they have three or four displayed items (figures and tables).

● 'Matters Arising' permits occasional short discussion of papers that have previously appeared in *Nature*. There is a limit of 300 words.

Manuscripts may be submitted either to London or Washington. Three typed copies should be submitted, each including lettered copies of figures. Typing (including references) should be double spaced. The title should be brief and informative. Pages should be numbered. References, tables and figure legends should start on separate pages. Experimental detail vital to the paper yet which would interrupt the narrative is best placed in the figure legends. Units should conform to the Système International. Greek characters should be identified in the margin on their first appearance. Equations should occupy single lines. If possible, exp (a) is preferred to e^a if 'a' is more than one character. Articles should be accompanied by an abstract of not more than fifty words listing the main conclusions

References are indicated by superscripts in the text. See any contemporary *Nature* for style, but note:

(i) To refer to several references by the same author at once, only one reference number need be given.

(ii) Cite first and last pages.

Abbreviations should follow the *World List of Scientific Periodicals*, fourth ed. (Butterworth, 1963–65). Symposia are often difficult to refer to and only published or soon-to-be-published volumes should be mentioned in references. Their publisher and place of publication should be clearly indicated. 'Personal communication' and 'unpublished work' should be incorporated in the text.

Artwork should be sent with the manuscript and clearly marked with author's name and the figure number. Line drawings to be in Indian ink on heavy cartridge paper, tracing paper or similar materials: thin, shiny, torn or folded material should be avoided. Scales for graphs should be chosen such that simple figures will not occupy an inordinate amount of space. Most figures are reduced to one column width so originals should be about as wide as a page of *Nature*. To enable figures, particularly maps, to be edited in the same style as the text they should contain only essential material. Ideally, an unlabelled original and a labelled copy should be provided: labelling on photographs should, if possible, be avoided entirely. Records, charts etc. which will not reproduce well should be redrawn or simply described in the text. We are always glad to see artwork for possible use on the cover, but cannot guarantee its return.

In order to save on postal expenses we return only the top copy and artwork of manuscripts that we cannot publish.

A fuller guide appeared in *Nature* (246, 238; 1973).

Two members of *Nature's* staff leave early in October.

Roger Woodham, Deputy Editor, is being promoted to Publishing Manager of Macmillan Journals Ltd. He has been with *Nature* for six years, during which time he has turned his hand to almost everything in the journal, but most recently to the physical science manuscripts and to maintaining the quality of production.

Allan Piper is leaving to become a writer with *The Oilman*. He has been with *Nature* for three years during which he has been a sub-editor, book-reviews editor and, for the last year, Assistant News Editor.

Cover picture

Stonehenge.
Photograph by Richard Gee.
[Barnaby's]

See pages 465 and 533.

OVER AND OVER AND OVER AND OVER AND OVER AND OVER AGAIN

Cover picture

The hydrogen-rich southern galaxies NGC 1510 and 1512. Photographed at the Anglo-Australian Observatory by B. A. Peterson.
On page 573, M J. Disney discusses the visibility of galaxies.

Volume 263   October 14, 1976

M

Editor
David Davies

Editorial Staff
Martin Goldman   Miranda Robertson
*Sandy Grimwade   Fiona Selkirk
Eleanor Lawrence   Stuart Sharrock
Mary Lindley   Chris Sherwell
Peter Milford   Hilary Taphouse
Peter Newmark   Robert Vickers
Colin Norman   Mary Wade
*Washington office

Publishing Director
Jenny Hughes

Display advertisement enquiries to:
Richard Webb in the London Office
or to
Mary Wade in the Washington Office
Telex: 64280

Classified advertisement enquiries to:
Advertising Department,
Nature,
4 Little Essex St
London WC2R 3LF
Telephone (01) 240 1101
Telex. 262024

Subscription enquiries to:
Macmillan Journals Ltd, Brunel Road,
Basingstoke, Hants. RG21 2XS
Telephone: Basingstoke 29242

Price

	UK	£35
Air/surface	USA and Canada	US$98
Surface	Rest of world	US$82 50
Airmail	{ Europe	US$94
	{ Rest of world	US$130

Cover picture
by Brian Grimwood.
Christopher Longuet-Higgins describes
(page 646) a computer program to
transcribe a live musical performance
into standard notation.

Vol. 263 No. 5579   October 21, 1976

Ready for rapid separation

Pharmacia prepacked, disposable Columns PD-2 and PD-10. Prepacked with Sephadex G-25 Medium for maximum economy and convenience in routine laboratory use.

Columns PD-2
Bed volume, 1.5ml

Suggested applications include: analysis for total serum thyroxine, analysis for serum triiodothyronine, removal of excess reagents from labelled proteins and removal of interfering ions before enzyme determinations.

Columns PD-10
Bed volume, 9.1ml

Desalting and buffer exchange with sample volumes of up to 2.5 ml in less than five minutes.

Pharmacia
Fine Chemicals

Cover picture
The eye of the shrimp.
See page 764.

Genetic guidelines: handle with care

PERFECT democracy or perfect dictatorship? That is a question which might well have to be asked now that the *Report of the Working Party on the Practice of Genetic Manipulation* has been published (Cmnd 6600, HMSO, 50p). Its recommendations embody a principle that Britain already knows in respect of its pay policy: that, threatened by a legislative hammer, everyone 'voluntarily' acquiesces in publicly announced guidelines.

The working party divides experiments into four categories (depending on their hazards) and provides a code of practice to cover each category (for details, see page 4). Researchers will submit details of their experiments to a Genetic Manipulation Advisory Group (GMAG), which will vet and categorise the work. They can then decide voluntarily whether or not to abide by the advice of the GMAG and to follow the code of practice.

But at the same time, details of the experiments will have to be submitted to the Health and Safety Commission (HSC), which has turned its eye specifically to all forms of genetic manipulation of micro-organisms, including those which preceded the advent of the new technology for recombinant DNA. The HSC will, through its inspectors, police the work using the existing (and very wide ranging) Health and Safety at Work Act, which it administers. The great advantage of this Act is that it applies to all employers, in industry, public laboratories and universities. This has allowed the working party to avoid the problem being faced in the United States, where guidelines control only research sponsored by the National Institute of Health (NIH) and leave industrial and defence-related work largely untouched.

So unspecific is the legislation, in fact, that it could probably cover any laboratory which carried out any experiment deemed to be hazardous to those who worked there. But, the HSC is now considering what special provisions it might need to give it more specific and authoritative control of genetic manipulation experiments. So the Act will be reinforced by a set of regulations, a draft of which has been produced. Their final form will depend on suggestions and recommendations received between now and November. One possibility is that the HSC will adopt the commendably detailed Code of Practice proposed by the working party, thus giving it a basis in law.

The GMAG, which will produce an annual report, will clearly be a powerful body. To its credit the working party offers the GMAG flexibility by refusing to go into the kind of detailed categorisation of experiments produced by the NIH in its analogous regulations; the preference, plainly, is for learning by experience. In addition the working party has considered not just the medical hazards of genetic manipulation in isolation, but also the potential dangers to plant and animal populations, which are included within the GMAG's scope.

The members of the GMAG, which will be responsible to the Secretary of State for Education and Science, will be announced as soon as possible. The main problem is likely to be the initial bottleneck in dealing with the batch of proposed experiments, some of which have lain dormant for two years. The GMAG is going to have to work speedily to overcome its own inexperience and the growing impatience of those now being encouraged by the working party, who have postponed their most exciting experiments.

Even more, of a problem may face the biological safety officer, whom the working party recommends be appointed in every laboratory carrying out genetic manipulation experiments. His task, particularly in the higher-risk laboratories, would be as onerous as it is important. And in the event of any prosecution under the Health and Safety at Work Act, he could not necessarily pass the buck on to the head of department to whom he is responsible. Given this, and the disinclination of most people to control the work of their colleagues, the job of biological safety officer may not prove to be a popular one. Might there not be a case, particularly in those laboratories with a major interest in genetic manipulation, for the appointment of a suitably trained person to that specific and sole role?

There is one slightly worrying aspect about the way the proposals have emerged. It relates to the British distaste for debate in full public view. Thus far, it has certainly been possible to state a view and have it considered by those deliberating the issues—but there has been nothing like the openness afforded in the United States. Greater openness would have produced an equally satisfactory result, and perhaps as quickly. But it would have also have meant greater legitimacy for the final course of action. The effectiveness of the proposals will be a reflection of how well or badly British science manages itself.

Off the tobacco road

UK cigarette manufacturers have reached
a crucial stage in the development of
substitute tobacco. **Allan Piper** reports

CIGARETTES containing what have
become known as "supplements" could
be on the UK market within 6 to 9
months. Rigorous health trials, recom-
mended by the UK Independent Scien-
tific Committee on Smoking and
Health, should be completed soon, and
three major tobacco companies with
British interests—Carreras Rothmans,
Gallahers and Imperial Tobacco—are
then expected to submit their results to
the committee.

The committee is an advisory body to
the Department of Health and Social
Security (DHSS) set up in 1973 under
the chairmanship of the Vice-Chancellor
of Birmingham University, Dr R. B.
Hunter. If it is satisfied that the find-
ings meet the guidelines it laid down
last year, the committee is expected to
recommend the licensing of supple-
ments for immediate commercial
development in Britain.

One company won't be on the early
bandwagon. Courtaulds decided last
month to hold back on its development
of a tobacco substitute. The decision,
however, was merely the commercial
judgment of a company temporarily
sidestepping an area of heavy pioneer-
ing expenditure which is already
covered by the others in the field.
Once supplements have been broadly
sanctioned, of course, Courtaulds can
jump straight back into the arena,
reasonably certain that the £3 million,
which it needs to carry its own product,
Tabrelle, through the Hunter tests, is
unlikely to disappear down the drain
The company will be up to two years
behind those already through the trials.
But it can take this course because
Tabrelle is essentially similar to the
other products being tested by the
tobacco companies.

Two types

There are, in fact, only two basic
types of supplement. One is made up of
largely inorganic materials called inert
diluents, the other is based on organic
cellulose derivatives. Both use an
organic binder. On current medical
reckoning—that the cigarette cancer
risk stems directly from the tar con-
tent of smoke—both kinds of supple-
ment provide a safer alternative to
ordinary tobacco. Organically based
varieties, such as Tabrelle and the un-
imaginatively named NSM ('New
Smoking Material') give less than

40% of the amount of tar produced by
ordinary tobacco. With Cytrel, the
main inorganic variety, even that
reduction can be improved two-fold.
This relative advantage is reinforced
further because the organic NSM pro-
duces more carbon monoxide, alde-
hydes and so-called furam derivatives
than tobacco, perhaps posing an addi-
tional, though as yet undetermined,
health threat.

In the event, none of the major
tobacco companies in Britain has
plumped squarely for one type of
supplement rather than the other.
Gallahers is maintaining an equal
interest in both Cytrel and NSM. And
while Carreras Rothman and Imperial
Tobacco have moved more positively in
one particular direction, respectively
favouring Cytrel and NSM, both com-
panies have kept their eyes on the
alternative option.

These preferences have resulted in
close research links between the retail-
ing tobacco companies and the manu-
facturing chemical corporations.
Carreras Rothman have an arrange-
ment with Celanese, the US makers of
Cytrel, while the tie between Imperial
Tobacco and ICI has led to the forma-
tion of the joint consortium NSM Ltd,
planned as an eventual commercial
manufacturer and supplier. In Ger-
many, the chemical giant Bayer has
just terminated a liaison with
Reemstma, following the spectacular
failure there last year of two brands
containing supplements. As for co-
operation betwen the companies and
the Hunter Committee, this is coordi-
nated by the Tobacco Advisory Council
and has apparently been effective, in
spite of misgivings within the industry
about the Committee's strong academic
make-up.

Three stages

The Hunter guidelines separate tests
into three stages, with an interim
check at the end of Stage I. This first
stage, already passed safely, included
comprehensive studies of smoke chem-
istry, together with inhalation tests
using rats and monkeys over a period
of at least six weeks. Stage II involved
short-term clinical studies of respiratory
irritation in human volunteers. At this
stage the tobacco companies were also
free under the guidelines to conduct

controlled consumer trials to assess
such characteristics as flavour accept-
ability.

The third stage, now virtually com-
pleted, involves the most rigorous tests
of all. Studies at this level fall into
three categories: carcinogenicity trials,
which include the classic mouse skin-
painting test; reproduction and terato-
logical studies, designed to assess
possible side-effects of smoking supple-
ments; and, finally, life-long inhalation
studies on rats and other species, which
in Britain has already led to the con-
troversy over the use of baboons and
beagles.

In a world context the Hunter
guidelines are unique, and most
interested corporations outside Britain
hope that once their UK competitors
have borne the expense of breaking
new ground, Hunter Committee ap-
proval can be used to ward off future
criticism of supplements on respective
domestic fronts. The £3 million cost
for the tests carry total development
costs for the tobacco companies in
Britain to around £10 million each over
the past five or six years.

When the committee licences the
supplements, marketing will begin. For
historical tax reasons they are still
grouped with additives in the Finance
Act, but new legislation to bring them
under the Medicine Act is expected to
be through parliament within the first
couple of months of next year. Without
hitches, cigarettes containing up to 25%
supplement could thus be on the British
market by March 1977.

NSM Ltd already has a large pro-
duction plant ready at Ardeer in Scot-
land, a £14 million investment by
Imperial Tobacco. Reports indicate
that Courtaulds could have Tabrelle in
commercial production within a year of
the Hunter green light for an additional
investment of £10–20 million. And
from across the Atlantic, Celanese can
be expected to provide all the Cytrel
needed in the early stages of UK
marketing. Similar products such as the
Anglo-American Batflake and Bayer's
still mysterious blend can also be ex-
pected to emerge rapidly once they
have passed through the Hunter tests.

Failure to win licences at the first
attempt would mean a major setback,
but the search for a safer cigarette will
certainly not be abandoned, even
though supplements are by no means
universally accepted as the most effec-
tive way to reduce the smoking risk.
Improvements to existing filters, parti-
cularly through ventilation, offer
another possibility, as does an increased
porosity of cigarette papers. The search
meanwhile continues for that elusive
premium blend of tobacco which gives
maximum possible taste with minimum
possible tar. ☐

The OTA on the EPA

Wil Lepkowski looks at the US Environmental
Protection Agency, the subject of a recent report.

WHEN the Environmental Protection
Agency was formed in late 1970, one
of the justifications for its creation was
that a single agency could take better
responsibility for restoring environ-
mental quality than a group of
separate fiefdoms. EPA, the argument
went, would view the environment as
a single, interrelated whole through an
administrative structure that was itself
correspondingly integrated. Environ-
mentally conscious America heaved a
sigh of assurance that at last the quality
of life and the preservation of Space-
ship Earth were given the policy
enthronement they merited.

While the human cell may be the
perfect blending of structure and func-
tion, EPA decidedly is not. It has faced
many reorganisations of its research
and development branches, internal
controversies, budgetary disappoint-
ments and manpower freezes; it has
suffered from conflicting Congressional
mandates, court setbacks to its enforce-
ment decisions, political shoving mat-
ches, and has had its programmes
skewered by new energy policies. Little
wonder, then, that EPA continues to
find the exercise of leadership elusive.

In the past, EPA was buffeted by
policy fashion as well as by legitimate
currents of change. Until the late
1960s, for example, environmental
control was seen as a problem of
public health protection, and was thus
housed in the Department of Health,
Education, and Welfare (HEW).
Gradually, however, the emphasis
began to shift from defining health
problems as guidelines for environ-
mental control to achieving techno-
logical solutions. The health problem
obviously remained, but Congress
decided that HEW was no place for
the development of technology or the
understanding of complex ecological
systems, and EPA was created.

It can be fairly stated that EPA
never could boast of a consistent
research policy. Hopes ran high during
its first two years when systems
analysts descended upon headquarters.
But systems mythology could only but
clash with traditions rising out of wet
chemistry and sanitary engineering.
Goals seldom squared with realities,
especially in environmental regulation
where research strategies were usually
dictated by crises emerging each week.
To be an EPA researcher was to be
unhappy. Besides being bureaucratically
overarching but arthritic, the agency

was acutely politicised by the Nixon
Administration.

Five year plan

Out of the wreckage—but in many
cases reflective of it—rises a five-year
research and development (R&D) plan
that Congress's Office of Technology
Assessment (OTA) has just finished
analysing for the committee on science
and technology. OTA has little good to
say of it; indeed, during the assessment
process, EPA scrapped the word "plan"
and substituted "outlook". It says in
essence that EPA's plan is not a plan
and as such is no guideline for helping
direct long term policies for the
country.

"Foremost among the shortcomings
in the R&D plan," OTA's report says,
"is EPA's failure to indicate a com-
mitment to long range research and,
as a corollary, an excessive focus on
short term R&D issues related directly
to the enforcement and/or achievement
of EPA's current regulations. Accord-
ingly," it goes on,

. . the Plan emphasises the development
and demonstration of control technologies.
In many cases, however, the larger prob-
lems involve social, economic and institu-
tional patterns which not only impede
technical solutions but which require
nontechnical approaches. To develop
effective overall environmental manage-
ment strategies will require more system-
atic and sustained socioeconomic research
efforts than those specified in the Plan.
An added R&D emphasis on long range
environmental concerns and a more res-
ponsive role to its lines of responsibility as
coordinator of Federal environmental
R&D would do much to enhance EPA's
effectiveness and credibility.

The US is obviously already envi-
ronmentally conscious. Whether it is
environmentally cleaner is another
question. The recent outbreak of the
so-called "legionaires disease" in Phila-
delphia appears to be caused by a toxic
metal whose origins and precise iden-
tity are still under investigation.
Evidence points to nickel carbonyl
from microcapsulised paper that re-
quires no carbon paper for the typing
of copies. Whatever the cause, the
point is that toxic chemicals are every-
where in the environment, and EPA
must always play a "catch-up" role.

But anticipation is the name of the
environmental game in its ideal, and
EPA's record as an anticipatory re-
search agency is a sorry one. Yet, anti-
cipatory research is a sorry subject
given the complexities of industrialisa-
tion and ecosystems. The obvious alter-

native is to require industry to stop
producing anything toxic to begin
with. EPA would administer the pend-
ing toxic substances control Act which
would do just that, but Congress seems
hesitant to pass a law so distasteful to
industry. It has been "pending" for
almost a decade. If passed, however,
EPA would become in theory much
like the Food and Drug Administra-
tion with its main regulatory goal being
prevention rather than reaction. Won-
dering what to react to has always
produced bureaucratic psychosis.

EPA's problems over the last few
years have been money, manpower and
morale. The EPA report doesn't
appraise the morale problem because
the charge was only to assess the Plan
itself. But the panels which gathered
to make the critique were composed
of individuals familiar with the agency's
internal problems and the moribund
psychology of the agency. That issue
wasn't discussed, despite its centrality.
The research arm of the agency is
plainly dejected. Individual scientists
know what needs to be done to begin
anticipating the environmental future,
but they know too that doing it would
require ten times the present research
budget of $250 million.

Questionable help

Thinking about environmental futures
is often a pathway to dejection, especi-
ally considering the fact that indus-
trialisation of the less developed
nations entails no pollution control to
speak of. And oceanographers have
long since despaired to any effective
pollution control policy for the oceans.

How much of a help OTA's critique
of the plan will be is questionable. It
says it isn't a plan. It says the main
issues are socioeconomic and EPA
hardly knows how to ask the right
questions in such a fuzzy area of para-
science. It says EPA knows next to
nothing about the long term effects of
pollution. It says EPA has a weak
health research capability. It is a
litany of deficiencies. The sorry fact is
that the same criticisms were made of
environmental research 15 years ago.

The findings have long since been
seen and mulled over by the House
science and technology committee. But
the practical and philosophical defi-
ciencies pointed out in the OTA report
transcend solution without profound
changes in consciousness among those
who make budgets and plan strategies.
And, of course, the entire milieu
within EPA would change once Con-
gress passes the toxic substances bill.
But that would mean further growth
in the regulatory bureaucracy and, of
course, larger costs of doing business.
The present Administration opposes
both.

BRITAIN_____

Genetic manipulation: guidelines out

Eleanor Lawrence studies the report on genetic manipulation experiments and the consultative document on regulations to control them.

GENETIC manipulation experiments involving artificially prepared recombinant DNA elements are to go ahead in the United Kingdom under certain specified conditions. Those intending to do such experiments must notify the Health and Safety Commission (HSC), an arm of the Department of Employment, and will also notify an independent Genetic Manipulation Advisory Group (GMAG) which will advise them of the conditions of containment required. The HSC is seeking powers to enforce the conditions recommended by the GMAG, and has put forward a set of draft regulations.

The voluntary code of practice and the proposed establishment of a national Genetic Manipulation Advisory Group are the main recommendations of the long-awaited report of the Williams Working Party on the Practice of Genetic Manipulation, published last week with government approval*. The aim of the technical guidelines and the machinery to put them into practice is to protect workers and the general public from the possible hazards presented by bacteria into which foreign genes have been transplanted by novel techniques developed several years ago in the United States.

These techniques enable DNA from another source (be it human, animal, plant, virus or an unrelated bacterium) to be spliced on to DNA from a phage or plasmid vector and introduced into a bacterium. The fear is that bacteria modified in this way (especially the common gut bacterium *E. coli*, a favourite of the genetic engineers) might present unpredictable dangers because of the new genes they possess, especially if the inserted DNA is not properly characterised or is known to contain potentially dangerous genes. The dangers could apply not only to the human and animal populations: harmless soil or plant bacteria into which genes from plant disease agents have been introduced could possibly pose a threat to crops.

But the advantages, both scientific and practical, which may eventually accrue from these new techniques are so great that there has been enormous pressure for the work to go on. Not

* *Report of the Working Party on the Practice of Genetic Manipulation* (Cmnd. 6600, HMSO, 50p).

only does it make it possible to clone particular genes and thus get to grips with many tricky problems in genetics, but specially engineered bacteria might eventually be induced to express the inserted genes and act as custom-built mini-factories making a valuable product quickly and cheaply.

Code of practice

The Williams working party was set up in response to recommendations contained in the Ashby Report which, in 1975, first considered the implications of the new genetic engineering techniques. Its brief was to establish a code of practice for conducting these experiments as safely as possible, and to define the role of the central advisory group which Ashby proposed. This, it was envisaged, would be responsible for ensuring that workers intending to perform genetic manipulation experiments were fully conversant with the safe microbiological techniques required.

Addressing the issue of matching particular experiments to the appropriate containment level, the working party has distinguished some 20 separate types of genetic manipulation experiment (see table) which they have then assigned to one of four levels of physical containment.

● **Category I**: Experiments must be carried out in conditions suitable for containing common human pathogens. (Because the introduction of alien DNA into a bacterium might have unforeseen consequences, the working party recommends that these precautions must be taken even when the source DNA, the vector and the host bacterium are all perfectly harmless in themselves).

● **Category II**: As I, but in addition the laboratory must be sited away from areas used by the general public and must have controlled air flow and an exhaust-protective cabinet for aerosol-producing operations.

● **Category III**: As II, but in addition access to the laboratory must be through an airlock only (which should contain washing facilities), all effluent from the laboratory must be decontaminated, and there should be an autoclave inside the laboratory.

● **Category IV**: Conditions should be equivalent to those already recommended in the Report of the Working Party on the Laboratory Use of Dangerous Pathogens for the most dangerous human and animal pathogens. (There are only a few such laboratories

in existence at present, for example those at the Center for Disease Control at Atlanta, Georgia, and at the Microbiological Research Establishment at Porton Down in the UK).

Genetic manipulation experiments can also be made safer using host bacteria and vectors which have been disabled so that they cannot grow outside a particular laboratory environment. Hosts can be crippled by introducing a requirement for some particular growth factor not found in the human or animal body, and disabled vectors which can only infect a special laboratory strain of host can be made.

The working party acknowledges the value of disabled organisms by assigning a lower containment level to experiments in which they are used, and recommends that research to develop new and better disabled strains be encouraged and the strains be made freely available.

Recommendations for implementation

The working party's recommendations are also concerned with the machinery needed to ensure that the code of practice is in fact followed. It recommends supervision at two levels, nationally

---

What is suggested ?

The Williams report recommends that:
1. experiments in genetic manipulation, conducted in appropriate conditions of physical and biological containment, should be encouraged.
2. further work should be done on the development and characterisation of disabled organisms and that any which are developed should be made freely available to all workers in the field.
3. no genetic manipulation experiment should be undertaken in containment conditions less stringent than those used for work with common pathogens.
4. the code of practice (in Appendix II) should be adopted as a basis for the conduct of these experiments.
5. every laboratory conducting these experiments should have a safety committee and a Biological Safety Officer.
6. appropriate training should be made available and be required for all research workers, technicians and biological safety officers in genetic manipulation laboratories.
7. a Genetic Manipulation Advisory Group (GMAG) should be established to advise on appropriate precautions for the conduct of these experiments.
8. the GMAG should be separate from the Dangerous Pathogens Advisory Group (DPAG) although there should be liaison between the two groups.
9. a system of voluntary control should be established as quickly as possible.
10. regulations should be made under the Health and Safety at Work Act to require laboratories to submit experimental protocols to the GMAG.

through the GMAG and, on the laboratory floor, in the person of a Biological Safety Officer, to be appointed in every laboratory in which genetic manipulation experiments are being carried out.

The Williams working party sees the central advisory group as crucial to the success of any control measures. One of its main functions will be to advise on the category into which a particular experiment should fall and on whether the laboratory is competent to carry it out. To do this the GMAG will need to maintain records of the facilities available in different laboratories; it should in time establish a register of approved laboratories. It should also review experimental protocols regularly and modify the code of practice in the light of new developments in biological and physical containment, and advise on training. The GMAG will be set up in the near future with a secretariat within the MRC, who will be responsible to the Secretary of State for Education and Science.

The biological safety officer will be the key person in the enforcement of correct procedures at the individual laboratory level. It will be his job to familiarise himself with the required techniques, and check that experiments are being carried out correctly; he will be responsible for training staff, investigating accidents and maintaining the security of the laboratory.

Recognising the "restraint already shown by the scientific community" since the voluntary moratorium on this work called for by American scientists two years ago, the working party now recommends that certain experiments which they consider do not present a serious hazard (Categories I and II, see table) should go ahead immediately, subject to notification to the GMAG. Experiments which fall into categories III and IV should not proceed until proposals have been approved by the GMAG.

In effect, the working party's technical guidelines are very similar to those published in July by the US National Institutes of Health, though there are important differences between the American and British approach. The US code of practice applies only to work carried out with NIH grants; the UK recommendations will apply to university government and industrial laboratories. The working party has also deliberately avoided the precise assignment of experiments with named organisms to particular categories.

Statutory control

The Working Party recommends that in the first instance the controls should be voluntarily imposed. But the draft regulations proposed last week by the Health and Safety Commission would provide some measure of statutory control without special legislation having to be passed.

The Health and Safety Commission is a government agency under the aegis of the Department of Employment set up to administer the Health and Safety at Work Act. This Act provides very wide powers to bring in regulations covering any work which might pose a hazard to the health and safety of work people and/or the general public.

The draft regulations, which are now being circulated as a consultative document to interested parties for comment, provide for compulsory notification to the Health and Safety Executive of all experiments which are "intended to alter, or likely to alter, the genetic constitution of any microorganism . . ." As it stands this blanket definition would cover practically all bacterial genetics experiments, and the Executive is asking the scientific community to recommend which conventional types of experiment should be exempted.

In addition the Executive wants to be able to enforce the precautions recommended by the GMAG through the Executive's Inspectors. They can already enforce general measures taken to avoid danger to health, but the Commission is inviting comments from any interested party on the need to include more specific provision for ensuring that the GMAG's advice is taken.

The ultimate sanction for researchers remains the witholding of a grant, however, and the working party recommends that, legal enforcement or no, government departments, research councils and other grant-giving bodies make grants conditional upon investigators following the advice of the GMAG. The working party also hopes that scientific papers will include a statement of both the physical and biological containment procedures used. □

Suggested categorisations for some typical experiments			
Source of nucleic acid	Specification of nucleic acid sequence	Vector/Host System	Category
Mammals	Random	Phage or plasmid/bacteria, not disabled	IV
	Random	Phage or plasmid/bacteria, disabled	III
	Purified*	Phage or plasmid/bacteria, not disabled	III
	Purified*	Phage or plasmid/bacteria, disabled	II
Amphibians and reptiles	Random	Phage or plasmid/bacteria, not disabled	III
	Random	Phage or plasmid/bacteria, disabled	II
	Purified*	Phage or plasmid/bacteria, not disabled	II
	Purified*	Phage or plasmid/bacteria, disabled	I
Plants and invertebrates and lower eukaryotes	Random	Phage or plasmid/bacteria, not disabled	II
	Random	Phage or plasmid/bacteria, disabled	I
	Purified*	Phage or plasmid/bacteria, not disabled	I
Mammals Amphibians and reptiles Birds	Random	Virus capable of infecting man or growing in tissue culture cells	IV
	Purified*	Virus capable of infecting man or growing in tissue culture cells	III
Viruses pathogenic to vertebrates	Random	Phage or plasmid/bacteria, disabled	IV
	Purified*	Phage or plasmid/bacteria, disabled	III
Animal viruses, non-pathogenic to man	Random	Phage or plasmid/bacteria, disabled	II
Bacteria specifying toxins virulent to man	Random	Phage or plasmid/bacteria, disabled	IV
Plant pathogenic bacteria	Random	Phage or plasmid/bacteria, not disabled	II
Plant viruses	Random	Phage or plasmid/bacteria, not disabled	II
Bacteria or fungi non-pathogenic to man, animals or plants	Random	Phage or plasmid/bacteria, not disabled	I

* The term "purified" means fractions with little chance of including any unrecognised extraneous sequences. It is of course possible to have sequences selected because of their pathogenicity and these would raise the level of containment required.

Source: *Report of the Working Party on the Practice of Genetic Manipulation* (Cmnd. 6600, HMSO).

FRANCE

Musical chairs in science policy

As talk of changes at the highest levels of French government grew last month, the country's directors of research were already a few steps ahead. A Correspondent reports

THE directors of French scientific and technical research have recently been involuntary participants in a game of musical chairs organised by the government. First, the president of the Centre National d'Etudes Spatiales (CNES), Professor Maurice Lévy, was dismissed by the Minister of Industry and Research, M Michel d'Ornano, for making concessions to striking staff that his assistant had refused to grant in his absence. He has yet to find another post. On the same day, Professor Curien, head of the Department of Scientific and Technical Research (DGRST), replaced Professor Lévy. The move, from a department which coordinates and implements all research policies (except military projects) to a directly operational post, has drawn much comment.

A few days later, Professor Curien, who was director of the Centre National de la Recherche Scientifique (CNRS) until 1973, was replaced at the head of the DGRST by Professor Grégory, who had previously taken his place at the CNRS. Then, to complete these acrobatics, which incidentally only involved physicists, the director of physics at the CNRS, Professor Chabbal, became director-general of the CNRS at the beginning of August.

The government—and in particular M d'Ornano—has shown in this series of decisions an unexpected authority and haste at a time when the credibility of those in power is constantly weakening. The circumstances of the CNES affair are fairly clear. The shrewd and effective Professor Lévy was forced to leave because of the Minister's change of mood—a change which neither his colleagues nor his advisers, who saw the opportunity to bring about other changes, sought to influence. A review of the allocation of the space budget between the programmes of the European Space Agency and the French programme had been due since European agreement on the level of the Agency's activities. The new allocation brought about some changes in the CNES budget which were to make some of its staff redundant, particularly at Toulouse.

With France wanting to guarantee Europe's autonomy regarding applications satellites, especially telecommuni-cations (the last breath of the great Gaullist doctrine of independence), the CNES gave priority to completing the launcher Ariane. (The Ariane project's director, Yves Sillard, became director of the CNES and assistant to the President a few weeks ago.) Under the threat of redundancies, CNES workers went on strike. The CNES directors were unwilling to give in to the strikers' demands, and maintained their position for several days, only to be overruled in the end by their president, who took the side of the employees. The government, however, intervened.

The causes of the stir this provoked in the research policy establishment are more complex. Two main characteristics differentiate research in France from research in other countries. The arrangement in the universities, where there is an even split between research and teaching, is unsatisfactory; information about what is going on does not permit either good financial management or national coordination. In accordance with the recommendations of the VIIth Plan, the Secretary of State for the Universities, who is responsible for university research and the CNRS, has created a "Research Mission" directed by Professor Denisse, who was Professor Lévy's predecessor as president of the CNES. The purpose of this mission is to re-organise university research.

The other characteristic of French research is the endemic weakness of industrial and technical research. French industry does not spend enough money on research, as M d'Ornano recently pointed out in parliament. Moreover, public funds normally allocated to industrial research are often put to other uses financing industry. The DGRST, which is for the most part made up of specialists in basic research and is directed by a laboratory physicist, often has great difficulty understanding industrial problems. This results in disagreements with other administrative departments concerned with industry, technology and financial management.

Thus the DGRST often ends up giving its funds (FF640 million in 1976) to the same laboratories. In an attempt to resolve this problem M d'Ornano asked the fiery director of the Centre National d'Etudes des Telecommunications (CNET) for a report on industrial research last April. One of the conclusions of the report was that a sub-department of industrial research and technology should be set up, linked to the Department of Research.

Michel d'Ornano

This proposal was strongly attacked by the DGRST: by its staff on the one hand who, like any other body, do not like the idea of being restructured by a logic they do not understand; and on the other hand by the director, Professor Curien, who distrusts dyarchies. The director was therefore moved to the CNRS. He was replaced by Bernard Grégory, who is also a former director of CERN. The Minister meanwhile stated that a technological representative will be nominated.

The proposal was also strongly attacked initially by another department, the Department of Mines. Although not implied by its title, this department is concerned with industrial research. It is the fiefdom of "mining engineers", a technocratic elite formed by the greatest of the "Grandes Ecoles", the Ecole Polytechnique. The mining engineers have an undying sense of unity: some hold high positions in the Administration, others are responsible for a large number of French industrial enterprises.

With these connections and the technical control of enterprises and industrial products that have by law fallen to it, the Department of Mines has developed a very different concept of industrial research and development from classical French practice since 1958. It is interested in the traditional industries; it tries to encourage collective research, common to all firms in the same sector; it checks that each project follows the objectives of the government (in conservation of energy, improvement and quality of products, working conditions and so on) through a complex series of committees. In short, it can take action effectively on scientific and technical information, industrial ownership, standardisation—and on funds.

If the future technological representative does not share its philosophy, there will be serious conflicts. And the stakes are high, especially for those enterprises which draw extensively on the research budget. ☐

IN BRIEF

EEC uranium scheme

Last week's announcement by the EEC Energy Directorate of a £400,000 scheme to support uranium prospecting within the EEC member countries came just two months after the report from Euratom forecasting a shortfall in uranium supplies for Europe from the end of the decade. Euratom expressed concern that there was no coverage either by long term contracts or from known reserves of Community producers, and urged an acceleration in exploration and development. The

Energy Directorate says it hopes the new project will discover new uranium sources to secure the EEC's long term supply.

No postgraduate change

The UK Department of Education and Science has declared in a White Paper (Cmnd 6611, HMSO, 28p) that it does not intend to restructure the arrangements for supporting postgraduate education in Britain. It was responding, after a delay of 2½ years, to strident criticisms of the system

voiced by the House of Commons Expenditure Committee. The committee's main contention was that postgraduate education should be shaped chiefly by the needs of the country rather than by the demands of potential students. The DES, while agreeing in broad principle, says that bodies like the research councils are already sufficiently responsive to these needs. The DES also rejects the idea that postgraduate grants should be replaced by any kind of loan, because the short term savings in government expenditure would not be sufficient.

IN theory, the more data a scientist is able to collect, the more accurate his work should become. On the other hand, a great deal of information is expensively collected and then never properly used. Sometimes this is a symptom of oversupported research; a scientist who has to do his own field or laboratory work is generally careful how he uses his time, but if he can deploy a squad of assistants he may ask them to make many more observations just in case these may, eventually, prove valuable. And very occasionally they do. But most of the results of even the best planned research remain untouched in dusty archives or even end up in the waste paper basket, although some work is written up and submitted for publication before sufficient information has been collected. It is not easy to decide when is the right time to stop.

An even more difficult task for the research worker is to decide just what data should be collected. If, to obtain a grant or a fellowship, he has to draw up a detailed programme in advance, he is unlikely to make the same decisions as he would have made after he had become immersed in his problem, particularly if he were personally involved in the details of the work and did not delegate to assistants. It is therefore essential that all research projects should have some provision for review and redeployment at frequent intervals.

But where a team of investigators is involved, such changes of direction may be difficult, as all members may not agree on just what changes are desirable. Even the lone worker may be inhibited from modifying his programme when he realises that this is necessary, in case he upsets the members of some committee to which he reports. Committee members react in this way, and castigate those who suggest that their original submission was faulty, perhaps because they themselves feel guilty for not spotting

the faults earlier.

We may collect too many or too few data, and we may also amass figures which do not provide the information that was intended. This is unfortunately too common an occurrence, and one which may long

---

Testing rainwater
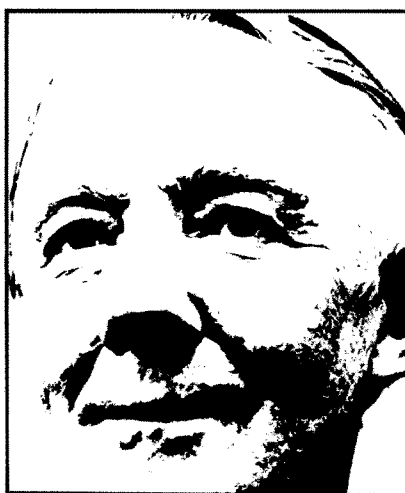
KENNETH MELLANBY

---

go unrecognised, as has recently been demonstrated in work on air pollution. Many people are studying problems which involve the contributions to plants and the soil made by the various chemicals contained in rain. Perhaps the best known is that of "acid rain" in Scandinavia, allegedly caused by pollutants, particularly SO_2, released into the air by British and European industry and carried north by the prevailing winds. Other work concerns the growth of crops and trees in polluted and clean localities. Unfortunately few of those involved are making systematic analyses showing exactly what substances are being brought down in the rain.

To many, including some of the investigators, this will appear to be a surprising statement, for numerous

people in many countries are analysing rainwater samples by the thousand, with increasingly accurate methods and more and more sophisticated equipment. There is little reason to doubt the accuracy of their measurements. Unfortunately they are not, for the most part, studying rain, but the liquid which they collect in their rain-gauges. This is something quite different, for it contains not only the chemicals which took part in the formation of raindrops, and those which were collected by the drops as they fell through the air, but also those which had been deposited, since the last shower, on the collecting funnel of the rain-gauge. The dry deposition on the gauge is generally greater, sometimes many times greater, than the quantity of the same chemical in the rain. If this is not realised, and measured, errors of a magnitude seldom experienced in research may arise.

Yet there is no excuse for this possible error. There is a little published information, available for some years, showing how much rainwater may differ from rain-gauge water. We cannot yet say exactly what the dry deposition of materials on any field, crop or forest will be, but there is much elegant work attempting these evaluations, and it is already clear that it greatly exceeds that in the rain, or even that in the already-enriched rain-gauge water.

Those who have spent time and effort analysing rain-gauge water, thinking they were analysing rain, need not despair. Their data are far from useless—they only fail to give the information expected. But this is a common phenomenon in science, and is one of the reasons to be suspicious of too highly orgaised investigations, particularly those commissioned by customers who think they can guarantee that the new information they seek can always be provided if only the cash is made available.

news and views

Stratified waters hold the key to the past

from M. Whitfield

THE history of the oceans and the atmosphere is, so far as we can tell, a story of the progressive oxidation of the primordial anoxic environment (F. T. Mackenzie, *Chemical Oceanography*, Riley, J. P. and Skirrow, G., eds, 1, ch. 5; 1975). Reducing conditions on the primitive Earth were an essential prerequisite for the formation of organic compounds from inorganic ingredients by a biological process. Oxygen was initially generated by photochemical processes in the atmosphere but large scale oxidation probably did not occur until photosynthetic blue-green algae began releasing free oxygen into the oceans some 3 billion years ago.

Over large areas of the present-day ocean, a permanent density discontinuity (pycnocline) arises as a consequence of the latitudinal variation in the intensity of incident radiation from the Sun which warms the tropical and subtropical surface waters and drives the cold polar waters into the ocean depths. This stratification stabilises the oceanic water column and restricts mixing between water in the sunlit euphotic zone and water from the abyssal depths. It is likely that the pycnocline has been an important feature of the structure of the oceans throughout most of their history (P. K. Weyl, *Science*, **161**, 158; 1968). After the development of photosynthetic cells the pycnocline would also mark the boundary between oxygenated and anoxic water and the resulting stratification would accelerate the build-up of oxygen in the upper layer and enhance the evasion of oxygen into the atmosphere. According to the Berkner–Marshall hypothesis (L. V. Berkner and L. C. Marshall, *J. atmos Sci.*, **22**, 225; 1965) the subsequent build-up of atmospheric oxygen played a crucial part in the development of higher life forms since it promoted the switch-over from a fermentative to a respiratory metabolism which was essential for the development of the nucleated eukaryotic cells that are the essential building blocks of all multi-

cellular organisms. The concomitant strengthening of the ozone layer provided a progressively more effective shield against ultraviolet radiation which made possible the colonisation of the energy-rich surface waters of the ocean and eventually the colonisation of the land. The interplay between processes that generate oxygen and processes that consume oxygen in stratified waters is therefore of central importance to any attempts to understand interactions between evolving life forms and their environment.

Pictures of a stratified Precambrian ocean previously developed to explain evolutionary processes (D. C. Rhoads, and J. W. Morse, *Lethaia*, **4**, 413; 1971) or geological phenomena (H. D. Holland, *Econ. Geol.*, **68**, 1169; 1973) have postulated a rather stable, layered system undergoing gradual oxidation over the past 2 or 3 billion years. E. T. Degens and P. Stoffers in this issue of *Nature* (page 22) relating the evolution of the early oceans to the recent history of present-day stratified waters suggest a more variable and dynamic history.

In strongly stratified systems the upper, oxygenated layer is relatively alkaline (*p*H 8 to 9), the levels of the major nutrients (nitrate, phosphate, silicate) are kept low by the incessant biological activity and carbonate minerals can be precipitated by inorganic or biological processes. The lower layer, by contrast, is more acid (*p*H 6.5 to 7) and contains hydrogen sulphide and abundant nutrients (ammonia, phosphate, silicate) and free carbon dioxide resulting from the decay of organic matter drifting down from the surface. This layer is also enriched in metals such as iron, manganese and nickel that are soluble in the reduced form. In the lower layer the variety of life forms is restricted to members of the sulphur biome (T. M. Fenchel and R. J. Riedl, *Mar. Biol.*, **7**, 255; 1970) probably the oldest ecosystem on earth.

The stability of the stratified condition depends on the dynamic balance

maintained between the water recirculation (recharging oxygen) and the degradation of organic matter (consuming oxygen). A rapid change in the depth of the pycnocline will cause a hiatus in the water chemistry which will leave its record in the sediments. A rapid downward movement will cause dramatic precipitations of iron, manganese and carbonate minerals and the corresponding injection of nutrients will stimulate plant growth in the oxidised layer. This growth will in turn increase the oxygen demand below the pycnocline and speed up the re-establishment of anoxic conditions. If the pycnocline now moves upwards some of the precipitated minerals (particularly the carbonates) will be re-dissolved and the incursion of hydrogen sulphide will ensure a massive extinction of bottom-living (benthic) organisms. The banded sediments of the Black Sea and the East African rift lakes observed by Degens and Stoffers, with the precipitates forming discrete layers rather than being dispersed amongst the regular detritus, suggests that the depth of the oxygenated layer has oscillated widely in response to climatic variations and to tectonic and erosional forces throughout the Holocene and Pleistocene. The pattern of variation is strikingly similar in the various locations and, furthermore bears a much closer resemblance to Cambrian and Precambrian sedimentary deposits than do modern oceanic sediments. These similarities suggest that the trend towards an oxygenated ocean might have suffered many reversals and that the details of these reversals are held in the sedimentary record, awaiting interpretation.

The oscillating pycnocline would certainly provide a particularly elegant explanation for the widespread banded-iron formations (rythmic sequences of iron-rich deposits interleaved with iron-poor cherts) that first appear in the sedimentary record 3 billion years ago and disappear fairly abruptly 1.2 billion years later. More dramatic upheavals of the system could also be engineered

to explain the periodic ecosystem collapses that have occurred at intervals of roughly 150 million years since Cambrian times (H. Tappan, *Palaeogeor.*, *Palaeoclimatol. Palaeoecol.*, **4**, 187; 1968). The analysis by Degens and Stoffers, particularly of deep sediment cores taken from the Black Sea by the Glomar Challenger, will enable such hypotheses to be tested quantitatively and should allow more definite statements to be made about the stresses that accompanied shifts of emphasis in the evolutionary process in the oceans.

These reminders of instability may also contain a word of warning. Natural control mechanisms are fairly delicately balanced and they often work by a process of overshoot and damping rather than by rigorous control within fine limits. As the Baltic countries are finding to their cost, increased oxygen demand resulting from industrial effluent can quickly generate oxygen-deficient bottom water in an enclosed sea. Similar problems have occurred in the Great Lakes of North America and have recently been reported in Lake Baikal. Furthermore, artificial upwellings proposed both as a source of power and as a means for stimulating planktonic growth in the surface waters of the oceans may also generate oxygen-deficient bottom water unless they are located with care.

Maintenance and evolution of repeated genes in eukaryotes

from D. Glover

THE genes for 5S RNA and ribosomal RNA are highly repetitive and occur as tandemly repeated clusters within eukaryotic genomes. These genes are separated by so-called spacer sequences which are not transcribed. Two main theories have been proposed to explain the mechanism by which tandemly arranged sequences are kept relatively homogeneous and how the genes and their spacers have evolved together. At the one extreme, are the sudden correction mechanisms such as Callan's "Master-Slave" hypothesis in which many genes are simultaneously corrected against a master template. At the other extreme, gradual correction mechanisms have been proposed which involve unequal crossing-over between homologous chromatids in such a way that variants could be spread or eliminated from tandem genes. The spacers of the genes for 5S DNA and rDNA in *Xenopus laevis* have now been subjected to elegant molecular analysis and this has cast some light on these problems.

The 5S DNA of *X. laevis* has an AT-rich spacer which consists largely of repeated units of simple sequence DNA (Brownlee *et al.*, *J. molec. Biol.*, **89**, 703; 1974). Carroll and Brown (*Cell*, **7**, 467; 1976) have recently described evidence which indicates that there is considerable heterogeneity with respect to length in this spacer sequence. Cleavage with the restriction endonuclease *Hind*III occurs at one site in each 5S gene and generates a series of closely spaced bands each of approximately 700 base pairs. These bands differ in length by regular increments of 14 base pairs which is the length of the simple sequence unit described by Brownlee *et al.* Carroll and

Brown went on to make partial *Hind*II digests of gradient purified 5S DNA, and have cloned these fragments. Each cloned DNA contains up to five of the 5S genes within a DNA plasmid of *E. coli* (*Cell*, **7**, 477; 1976). Denaturation mapping of these cloned 5S genes indicates that adjacent genes can be heterogeneous in length. It is unlikely that such arrangements could be found if sudden correction mechanisms were operating during evolution and Carroll and Brown come down strongly in favour of a mechanism of unequal crossing over between the AT-rich spacers in which the sub-repeats are essentially homologous.

Wellauer *et al* (*J. molec. Biol.*, **105**, 461; 1976) have now made a similar analysis of the ribosomal genes from *Xenopus laevis*. There are two classes of *Eco*RI fragments which can be excised from the gene unit. One of these is homogeneous for length whereas the other class of fragment is heterogeneous for length and contains non-transcribed spacer sequences. Four representatives of this latter class have been cloned within *E. coli* and then characterised by homoduplex and heteroduplex mapping. This has revealed two regions within the non-transcribed spacer which vary in length and seem also to consist of repetitious simple sequences, the subunit of which is shorter than 50 base pairs in length. In *X. laevis* the ribosomal genes of oocytes undergo substantial amplification, a process which is thought to occur by a rolling circle mechanism of DNA replication (Hourcade *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **70**, 2926; 1973; Rochaix *et al.*, *J. molec Biol.*, **87**, 473; 1976. If this is so, then one would predict that adjacent rDNA genes from oocytes

would be more homogeneous for length than adjacent rDNA genes from somatic tissue. Wellauer *et al.* (*J. molec. Biol.*, **105**, 487; 1976) have examined heteroduplex structures formed between either long single-strand chromosomal or amplified rDNA and a cloned segment of rDNA and have found that this is indeed the case. In a third paper, the same group has found that not only is there heterogeneity in spacer lengths, but also that the pattern of spacer lengths is significantly different from one animal to another so that it can be followed as a genetic marker (Reeder *et al.*, *J. molec. Biol.*, **105**, 507; 1976). The patterns of spacer heterogeneity are stably inherited from one generation to the next and only two animals out of 50 showed changes in their spacer patterns. The mechanism by which these two changes arose remains an open question.

The tandemly repeated gene clusters of *Drosophila melanogaster* are also producing their share of surprises. The nucleolus organisers of *D. melanogaster* are found on both the X and the Y chromosome. In wild-type males there is apparently no meiotic recombination and so the two nucleolus organisers are genetically isolated from one another on non-homologous chromosomes. The rRNA transcribed from either the X or the Y seems to be identical however. A single ribosomal gene unit from this organism has been described which from the evidence of DNA reassociation kinetics also contains about 3,000 base pairs of internally repetitious simple sequence. There are two major size classes of rDNA, one in which the repeating gene unit is 17 Kbases in length and the other in which the repeating gene unit is 11–12 Kb in length (Glover *et al.*, *Cell*, **5**, 149; 1975). Tartof and Dawid (on page 27 of this issue of *Nature*) have examined the distribution of these two units on the X and Y nucleolus organisers and find that whereas the 11–12 Kb unit is present on both the X and Y chromosomes, the 17 Kb unit is found only on the X. This suggests that indeed recombination is not occurring between the X and Y rDNAs, unless there is some specific mechanism for eliminating the 17 Kb units from the Y chromosome.

The mechanism whereby the identity of the rDNA sequences on the X and Y chromosomes is maintained is obscure, and indeed it will be interesting to have a comparison of spacer sequences on the 12 Kb and 17 Kb units which have been cloned in *E. coli*. The mechanism of the maintainance and co-evolution of these tandomly repeating genes may turn out to be even more complicated than the experiments with *Xenopus* 5S DNA and rDNA suggest.

T cell recognition at Cold Spring Harbor

from Martin Raff

The XLI Cold Spring Harbor Symposium on Quantitative Biology, entitled Origins of Lymphocyte Diversity, was held on June 1–8, 1976.

SINCE the symposium covered most of present-day immunology this is an attempt to cover just one area—how T cells recognise antigen. In my view, it was recent progress in this knotty area of immunology that provided much of the excitement at the meeting.

While it is clear that membrane-bound immunoglobulin (Ig) molecules serve as receptors for antigen on B lymphocytes, the nature of antigen-specific receptors on T cells has been elusive and, for the most part, unproductively controversial. Now, at last, half of the problem seems to be resolved, with the demonstration that at least one receptor on T cells is a new class of Ig heavy(H) chain, which may come to be known as "tau" (τ). The strongest evidence for this was provided by H. Binz and H. Wigzell (Uppsala University), who, following the lead of Ramseir and Lindenmann[1], have approached the problem using antibodies directed against antigenic determinants associated with the antigen-binding sites of Ig molecules—so-called idiotypic determinants.

They had previously shown that anti-idiotypic antibodies (raised in (Lewis × BN) F_1 rats against Lewis anti-BN alloantibodies, or against Lewis T lymphocytes) react with both T and B cells of Lewis rats which bear receptors specific for the major histocompatibility complex (MHC) antigens of BN rats[2]. Thus, T lymphocytes seem to express some of the same idiotypes as conventional antibodies—a notion supported by similar findings obtained independently by others, in rats[3] and mice[4]. Binz and Wigzell have now shown that whereas the idiotype-bearing (and BN-antigen-specific) receptors on Lewis B cells are 8S IgM molecules, those produced by T cells are polypeptide chains of molecular weight approximately 75,000 daltons, which may exist as dimers and which do not react with conventional anti-Ig antibodies or with antibodies directed against Lewis MHC antigens (which presumably include activity against both Ag-B and I-region-associated (Ia) antigens). Moreover, they reported that these in-

herited idiotypes are genetically linked to a Lewis Ig H chain allotype, indicating that the idiotype-bearing receptors on T cells are coded for in the H-chain variable region (VH) gene pool. Since the isolated T cell receptors seem not to be tightly associated with other polypeptides of different molecular weight, the simplest view is that single or dimeric Ig H chains, having a previously undescribed constant region, serve as receptors for antigen on T lymphocytes.

The findings reported by K. Eichmann (German Cancer Research Centre, Heidelberg) and K. Rajewsky (University of Cologne) and their colleagues were consistent with this view. They had previously used anti-idiotype antibodies raised in guinea pigs to show that both T helper cells and B cells in A strain mice use the same inherited VH genes for responding to Group A streptococcal carbohydrate[6]. They have now found that only those anti-idiotype antisera that recognise idiotypes genetically linked to H chain allotypes (and which are therefore related to the H chain alone) are able to stimulate (prime) both T and B cells *in vivo*, whereas those that react with idiotypes that require light (L) chains for their expression can only stimulate B cells In addition, using hapten-coupled nylon mesh to purify rabbit lymphocyte receptors, they reported the presence of VH allotype determinants on putative T cell receptors which bind antigen but fail to react with antibodies against constant region antigenic determinants of Ig L or H chains

The other side of the T cell recognition puzzle concerns the target antigens that T cells see. There is increasing evidence that T cells only respond maximally to conventional antigens when they see them in association with MHC antigens. Thus, in order for cytotoxic T cells to respond optimally to viral[6], haptenic[7] or minor histocompatibility[8,9] antigens on mouse target cell surfaces, they must see them in association with H-2D or H-2K gene products; T cells that proliferate in response to antigen-coated "macrophages" *in vitro*, recognise antigen in association with the Ia antigens of the "macrophage" (ref. 10 and W. Paul, National Institutes of Health); helper T cells seem to see antigen in association with Ia antigens (C. Pierce, Harvard Medical School; M. Feldmann and

P. Erb, University College London; H. Von Boehmer (Basel Institute for Immunology) and J. Sprent (University College London)), and T cells responsible for delayed hypersensitivity see some antigens (foreign gamma globulin, for example) in association with Ia antigens, and others (such as dinitrofluorobenzene) in association with H-2D or H-2K gene products (J. Miller, Walter and Eliza Hall Institute, Melbourne).

In most of these situations, the associative recognition of MHC antigens can involve syngeneic or allogeneic MHC determinants; however, the MHC context (syngeneic or allogeneic) in which antigen is first seen by T cells in a primary response is "remembered" by the primed T cells and must be the same MHC context if these cells are to give a secondary response to the antigen. This has been shown for cytotoxic T cells (Zinkernagel and Doherty; Von Boehmer and Sprent), helper T cells (Von Boehmer and Sprent; Pierce; D. Katz, Harvard Medical School) and T cells mediating delayed hypersensitivity (Miller), and explains most experimental results which had previously been considered examples of "syngeneic preference", where it was thought that T cells and macrophages[10], or T cells and B cells[11] had to be syngeneic for I-region determinants in order to collaborate. The meshing of the findings in all of these T cell systems is both comforting and exciting. Together, the observations underline the central importance of the MHC in T cell recognition, which must be one of the most important advances in immunology in recent years.

What does this mean in molecular terms? If the same inherited VH genes are used by helper T cells and B cells in A strain mice to recognise Group A streptococcal carbohydrate[6], and the helper T cells (but not B cells) recognise this antigen in association with Ia determinants, then it seems highly probable that the T cells use a different gene product for recognising the Ia antigens Since, so far, allelic exclusion has only been shown for B cells it is conceivable that two different VH gene products are involved, one maternal and one paternal. It seems more likely, however, that T cells make two genetically independent receptor molecules, just as B cells make genetically unlinked L and H chains; unlike the B cell receptors, however, the two T cell receptors may not be physically associated on the cell surface. More evidence for two independently coded T cell receptors is provided by Heber-Katz and Wilson[13] who found that T cells positively selected for reactivity against MHC antigens are as effective as unselected T cells in helping B cells respond to sheep erythrocytes. The

pressing need now is to identify the putative "second" T cell receptor; less pressing is the need to establish whether the "antigenic" and MHC determinants have to be physically associated, either on the target cell or "macrophage" surface, or as a shed complex[13]. In any event, the notion that T cells look primarily at a single compound determinant, in the form of antigen-modified MHC glycoproteins ("altered-self")[14] seems improbable.

If there are two classes of antigen-specific receptors on T cells, the undefined "second" receptor presumably recognises mainly self and foreign MHC antigens. Yet, Binz and Wigzell report[15] that 3–5% of Lewis T cells label with their anti-idiotype antibodies, which are thought to react primarily with those VH gene products that recognise BN MHC determinants. Therefore, it seems that both the VH and the putative "second" T cell receptor gene pools may be heavily committed to recognising MHC antigens and that Jerne[16] may have been right—yet again.

With the demonstration that Ia antigens are important in associative recognition by T cells, the list of I-region functions continues to grow. Various Ia antigens have now been identified on the surfaces of B cells and "macrophages" (see ref. 17), suppressor (D. Murphy and L. Herzenberg, Stanford; M. Feldmann, University College London) and helper (Murphy and Herzenberg) T cells, on enhancing and suppressing factors probably produced by macrophages (Feldmann and Erb) and T cells (T. Tada, Chiba University, Japan; and ref. 18), and on acceptor sites for such factors on T (Tada; Feldmann and Erb) and probably B lymphocytes[18]. The simplest view is that most, if not all I-region loci are involved with T cell responses or T cell regulation of B cell function, and that the I-region Immune response (Ir) genes, which regulate T-cell dependent immune responses to a variety of specific antigens, exist at many, if not all, of these different loci, and do not primarily code for T cell receptors. That immunity to a single antigen can be controlled by more than one Ir gene has already been established (B. Benacerraf *et al.*, Harvard Medical School; and refs 18–21) as has the fact that these genes can be expressed in macrophages (or in cells that copurify with macrophages)[18].

Until now, the biology of T lymphocytes has lagged well behind that of B cells—but the pendulum has reversed its swing. Unravelling the functions, biochemistry and genetics of the products of the MHC, now recognised to be central to T cell function, is providing most of the action in immunology these days and the Cold Spring

Harbor Symposium was a celebration for those involved.

1 Ramseer, H., and Lindenmann, J., *Eur. J Immun.*, 2, 109 (1972).
2 Binz, H., and Wigzell, H., *J. exp. Med.*, 142, 197, 1218, 1231 (1975).
3 McKearn, T. J., *Science*, 183, 94 (1974).
4 Eichmann, K., and Rajewsky, K., *Eur J Immun.*, 5, 661 (1975)
5 Hämmerling, G. J., *et al.*, *J. exp. Med.*, 143, 861 (1976).
6 Zinkernagel, R., and Doherty, P., *J. exp. Med.*, 141, 1426 (1975).
7 Shearer, G. M., *et al.*, *J. exp Med.*, 141, 1348 (1975).
8 Bevan, M. J., *Nature*, 256, 419 (1975).
9 Gordon, R. D., *et al.*, *J. exp. Med*, 142, 1023 (1975)
10 Shevach, E. M., and Rosenthal, A. S., *J. exp. Med.*, 138, 1213 (1973).
11 Katz, D., *et al.*, *J. exp Med.*, 141, 268 (1975).
12 Haber-Katz, E., and Wilson, D B, *J. exp. Med*, 193, 701 (1976).
13 Erb, P., *et al.*, *Eur. J. Immun.*, 6, 365 (1976)
14 Zinkernagel, R., and Doherty, P., *Nature*, 251, 547 (1974).
15 Binz, H., and Wigzell, H., *J exp. Med.*, 142, 1231 (1975)
16 Jerne, N. K., *Eur. J. Immun.*, 1, 1 (1971).
17 *Transplant. Rev.*, 30 (1976)
18 Munro, A. J., and Taussig, M. J., *Nature*, 256, 103 (1975).
19 Rude, E., and Gunther, E, *Progr. Immun.*, II, 2, 223 (1974).
20 Melchers, I, and Rajewsky, K, *Eur. J. Immun.*, 5, 753 (1975)
21 Zaleski, M., and Milgrom, F., *J Immun*, 110, 1238 (1973).

Finding a niche

from Peter D. Moore

THERE are few terms in ecological jargon which are used more frequently than the 'ecological niche' and there are few which have been used so loosely. This is particularly true in plant ecology where there have been few attempts to construct working definitions. The use made of the term by Elton (*Animal Ecology*, Sidgwick and Jackson, London; 1927) was essentially applicable to animals and referred to the role of a species in a community as far as its feeding and competitive interactions are concerned. In autotrophic plants the variety of functional roles available is obviously severely limited, so the Eltonian concept of niche has not proved particularly valuable.

Hutchinson (*The Ecological Theater and the Evolutionary Play*, Yale University Press, New Haven; 1965) has provided a concept of 'niche' which is of much greater value to plant scientists. If one considers all the possible environmental factors which influence a species, and the organism's response to these factors, then these can be regarded as axes in multi-dimensional space. Each species thus occupies a definable volume of 'hyperspace', which represents the set of conditions under which the species can survive. This potential or 'fundamental' niche, may not be identical to its 'realised' niche, for competitive interactions between species whose fundamental niches overlap may result in their restriction within their potential volume. Wuenscher (in *Handbook of Vegetation*

Science No. 6, edit. by Strain, B. R. and Billings, W. D., 39, Springer Verlag, Berlin, 1974) has discussed the Hutchinsonian concept of niche as it applies to plants and has also stressed how the concept goes beyond that of simple habitat. Any point within niche hyperspace can be described by the coordinates of all the environmental variables (habitat factors) and the response of the organism to these factors. The response element carries the niche concept beyond that of habitat alone.

In a community it is to be expected that the component species will occupy adjacent hyperspace volumes, since their basic habitat requirements will be similar In this case, the introduction of a new species can only be successful if other species are permanently displaced from at least part of their former niches. Total displacement will, of course, result in extinctions. An example of this situation was described in some detail, by Zaret and Paine (*Science*, 182, 449; 1973) following the introduction of an exotic cichlid fish, *Cichla ocellaris*, into Gatun Lake in the Panama Canal Zone. This predatory species caused considerable simplification in community structure by supplanting several native species. Examples from the plant kingdom are not as well documented, but they certainly exist. The abundance of the sycamore (*Acer pseudoplatanus*), introduced into British woodlands probably in the early sixteenth century, has almost certainly been achieved only at the expense of native species, but it would be difficult to prove this.

There are situations in which an introduced species appears to have no deleterious influence upon the native components of a community. If such a situation does occur, it implies that the niche of the invading species has no volume of overlap with those of the natives. It has recently been claimed by Hudson (*Brit. Birds*, 69, 132; 1976) that the North American ruddy duck (*Oxyura jamaicensis*) which has become feral in Britain over the past 30 years, has occupied a 'vacant niche' in that its ecological requirements overlap little with such native diving ducks as the pochard (*Aythya ferina*) and tufted duck (*A. fuligula*). In Hutchinsonian terms, the niche can only be defined in terms of environment and species response, so strictly one should not speak of 'vacant niches', however the idea of an untapped environmental resource is often referred to in this way. The former vacancy in hyperspace can only be recognised in retrospect.

Danin (*Oecologia*, 22, 251; 1976) has recently conducted a survey of annual plant species under desert conditions in the Dead Sea Valley of Israel. He showed that diversity falls with increas

MAN'S first picture of the surface of Venus caused great interest and some surprise as it revealed that Venus, like Mercury, Moon and Mars, had a cratered surface. The picture was obtained using simultaneous Doppler and time delay processing of 12.6 cm wavelength radar echoes reflected from the surface (Rumsey *et al.*, *Icarus*, 23, 1; 1974). Looking at a 1,500 km diameter circular area at the centre of the planetary disk on June 20, 1972, they found it to be nearly flat but cratered. The biggest crater was 160 km in diameter, the floor appearing to be on the same level as the surrounding terrain but with the rim 500 m above the floor. By contrast a lunar crater of this size would be about seven times deeper.

Venus and Earth have fairly similar masses and both seem to have large raised regions on their surfaces. The "continents" on Venus do not appear to be isostatically compensated so that active internal convection may be taking place coupled with a degree of surface volcanism. Craters will be formed both by volcanic activity and by meteoritic activity; however, the information needed to determine which phenomenon is more likely is still lacking.

Michael Tauber and Donn Kirk (NASA–Ames Research Center) have now attempted to determine the sizes of the stone and iron meteoroids which can penetrate the atmosphere of Venus and cause hypervelocity impact craters on the planet's surface (*Icarus*, 28, 351; 1976). Obviously the surface pressure of 90 atmospheres and the atmospheric density of 64 kg m^{-3} present a major obstacle to the passage of meteoroids Ninety percent of the meteoroids will have atmospheric entry velocities ranging from 10 to 40 km s^{-1} and if all of the kinetic energy lost by atmospheric

drag was available to heat and vapourise the meteoroid no meteoroids less than 10^{11} kg in mass would survive passage through the atmosphere Actually only a small fraction of this energy goes to heat the body, the great majority's being lost as it heats the atmospheric gases The complex interactions between the body and the high temperature–high pressure shock layer surrounding it make the mass

Venusian craters
from David W. Hughes

loss calculations very difficult. Tauber and Kirk find that bodies over 10^4 kg only lose a few percent of their mass by vapourisation, mainly because entry takes place so quickly. The velocity of the body can decrease drastically though.

A hypervelocity impact crater is produced when the shock pressure at the impact point exceeds the compressive dynamic strength of the target material. Venera 8 measurements suggest that Venus has a granite-like surface material and the authors calculate that basalt and iron projectiles with velocities greater than 1.1 km s^{-1} and 0.7 km s^{-1} will produce hypervelocity impacts and craters. For basalt and iron particles they find that it is only those larger than 100 m and 40 m that have final velocities above these limits. One complication is the possible mechanical break-up of the body before impact. This is caused by a combination of thermal stress from frictional heating and also air pressure acting on the high speed body as it is retarded in the atmosphere. Frictional heating is confined to a thin shell and probably does not affect large meteoroids. The atmospheric forces on the other hand become very large and may exceed the compressive strength. It seems

that for large bodies this only happens for a second or so before impact so there is not enough time for the cracks to spread throughout the body or for fragments to separate significantly So on Venus this effect does not change the size of the crater produced. The authors calculate that a stone meteoroid 1 km across, entering the Venusian atmosphere at 40 km s^{-1} will produce a crater 33 km in diameter. This diameter is proportional to the sixth power of the particle's density and its kinetic energy raised to the power 0.28. They also find that the smallest hypervelocity impact crater on the surface of Venus is about 150 to 300 m across. This contrasts dramatically with the Moon which has no atmosphere to retard or fragment the meteoroids and is pockmarked with craters of all sizes greater than a fraction of a micrometre

Venera 8 measurements (*Icarus*, 20, 407; 1973) indicate very low surface winds on Venus so the crater structure may be preserved for long periods of time. The radar picture taken by Rumsey *et al* had a resolution of about 12 km, so it cannot be used to test Tauber and Kirk's theory. Radar mapping from a planetary orbiter, or radar using the Arecibo radio telescope could provide the resolution In fact photography from a probe descending through the Venusian atmosphere is possible as the daylight illumination level at the surface is thought to be about 2% of that on Earth. The discovery of primary impact craters smaller than the 150–300 m calculated above would indicate that the impacts occurred when the atmosphere of Venus was less dense than it is now. So crater size distribution thus provides a straightforward indication of atmospheric evolution

---

ing soil salinity until eventually, in highly saline conditions, only one species, *Mesembryanthemum nodiflorum*, survives. Danin describes this gradient as one in which a physiological stress reduces the number of niches until, under extreme stress, a one niche habitat results. This situation does not preclude competitive interplay, but demonstrates the superior growth potential of the surviving species under the stress conditions In the Hutchinsonian sense it can thus be termed a one niche habitat What one must not discount, however, is the possibility that there could be unrecognised opportunities even here for species which have either not yet invaded, or may not even yet have evolved in the area. Could one term such an availability of

opportunity, such untapped resource potential — such unoccupied hyperspace, a vacant niche? □

Nuclear polysomes
from J. R. Tata

CELL biologists often emphasise the intracellular segregation of transcription and translation between the nucleus and cytoplasm as a major distinguishing feature between eukaryotes and prokaryotes Every so often, however, fresh controversy is generated by reports that challenge the validity of this intracellular compartmentation of macromolecular synthesis by demon-

strating that nuclei can also synthesise proteins (for a review, see Kuehl, in *The Cell Nucleus*, edit by Busch, H., 3, 345, Academic Press, New York, 1974). Many such claims have had a short life since these are based on indirect evidence, or fail to identify a special class of proteins synthesised within the nucleus. Perhaps the most important reason for the controversy over whether or not the nucleus possesses a protein synthetic apparatus may reside in the ease with which isolated nuclei are contaminated with cytoplasmic polyribosomes (Penman, *J molec Biol*, 17, 117; 1966), particularly those associated with the perinuclear endoplasmic reticulum (Lewis and Tata, *J Cell Sci*, 13, 447, 1973).

Jo Alene Goidl and coworkers at the

Roche Institute in New Jersey have now managed to demonstrate polyribosome-like structures from isolated HeLa cell nuclei meticulously freed of contaminating cytoplasmic polysomes by detergent treatment (Goidl *et al.*, *J. biol. Chem*, **250**, 9198; 1975). Exposure of these nuclei to polyanions (heparin, RNA and particularly poly (U)), which are known to disrupt nuclear structure (Coffey *et al.*, *Adv. Enzyme Regnt*, **12**, 219; 1974), caused the release of particles of 200–300 A in diameter. These unique particles were shown, in a heterologous translation system derived from Ehrlich ascites cells, to comprise mRNA or precursors of mRNA associated with ribosomes and the complex itself was shown to be able to carry out translation. Goidl and her colleagues also found that poly (U) used for the disruption of nuclei was not associated with these nuclear polysomes; yet they were capable of polypeptide chain elongation but not initiation, as judged by effects of inhibitors of cell-free protein synthesis. These findings may have some relevance to those of Bachellerie *et al.* (*Eur. J Biochem.*, **58**, 327; 1975), who characterised transcription complexes of ribosomes and particles containing HnRNA still bound to chromatin derived from nuclei of ascites hepatoma and Chinese hamster ovary cells. The question of whether such complexes synthesise proteins within the nucleus or whether they are "packages" for transport of mRNA into the cytoplasm, as Harris (*Nucleus and Cytoplasm*, Clarendon Press, Oxford; 1974) suggested, is still open. Be that as it may, in view of the past controversy about protein synthesis by isolated nuclei, Goidl *et al.*, quite prudently, still do not state categorically that their polyribosomal material is exclusively nuclear but imply that if these complexes are of cytoplasmic origin, then their release from the nucleus somehow requires disruption of chromatin.

In a wider context, two questions concerning the capability of protein synthesis by nuclei have now to be considered. First, if they do have this function, do they make special classes of proteins? There are periodic but often unconfirmed reports of such special proteins exclusively made in the nucleus, but it is almost certain that the major nuclear proteins (histones, acidic chromosomal proteins, DNA and RNA polymerases) are synthesised in the cytoplasm. Second, what is the likely magnitude on a cellular basis, of protein synthetic activity in the nucleus? By isolating detergent-washed nuclei from sea urchin blastula cells, labelled *in vivo* with ³H-leucine, Allen and Wilt (*Expl Cell Res.*, **97**, 151;

1975) concluded that no more than 0 2% of the total cellular nascent polypeptides could be recovered with the nuclei. Perhaps this figure may be higher for nuclei or other cells, such as thymus or nucleated erythrocytes, but it is still likely to be quite low. Whatever the extent of protein synthetic activity of the nucleus. Goidl and her colleagues have certainly reopened, and with convincing evidence, an old controversy. Past critics will certainly find new reasons to disbelieve translation of any RNA within the nucleus. One has only to go back fifteen years to recall the controversy raging around protein and nucleic acid synthesis in mitochondria to draw a useful lesson and keep an open mind on this unsolved problem. □

Estimating maximum quake magnitudes

from Peter J. Smith

ACCORDING to a well known empirical formula, the number (N) of earthquakes per unit time exceeding any magnitude (M) is given by $\log N = a + bM$. Both a and b are usually described as positive constants, although they are not universally constant; a varies greatly from region to region and b, though generally more restricted, varies with region, focal depth range and in some areas even with magnitude. Irrespective of these variations, however, implicit in the general relationship is the existence of an upper limit on earthquake magnitude. If a and b are taken to be the averages for worldwide earthquakes, this limit turns out to be about 9.0.

It would appear from the work of Perkins (*Earthquake Inf. Bull.*, pp 18–23, July 1972), and others that this maximum magnitude may be established on the basis of worldwide earthquake statistics covering the period of only the past 50–70 yr, during which time it may safely be presumed that at least one event of roughly maximum magnitude has actually occurred (for example, the magnitude 8.9 earthquakes of Colombia in 1906 and Japan in 1933). Unfortunately, the establishment of a global upper magnitude is largely of academic interest. A more common practical problem (for example, in the field of earthquake engineering) is the assessment of the likely maximum magnitude of earthquakes within a restricted local region where the seismic statistics are but a depleted subset of the global total.

In local areas where earthquakes are clearly associated with faults, a very rough estimate of likely maximum magnitude along a given fault may be

made from curves of fault length against magnitude based on recent seismic experience. Unfortunately, however, the correlation between earthquake magnitude and surface rupture is poor and varies with fault type. In regions where the faults are less clear, the prediction of the upper limit is even more uncertain; it can only be assumed that the earthquake of highest magnitude ever recorded in the area could be repeated anywhere within it The drawback of such a method is obvious An earthquake of the highest possible magnitude for any given zone may not have been recorded in the zone in historic time. Moreover, the period covered by local seismic statistics may be quite insufficient to define an expected maximum magnitude on the basis of the formula involving a and b.

But how long a sampling period would be adequate? According to S. W Smith (*Geophys. Res Lett*, 3, 351; 1976), if the Earth's major seismic zones with a total length of about 80.000 km produce several maximum magnitude earthquakes in 50 yr, then 20.000 yr of seismic monitoring would be needed to give comparable statistics on a local scale of, say, 100 km. Needless to say, such long term monitoring is out of the question. On the other hand, it is quite feasible in principle to obtain fault displacement over a comparable period from the geological record. What is required, then, is some way in which, in the absence of an adequate seismic record, the geological record may be made to yield an estimate of maximum seismic magnitude.

The essential link between geological data and earthquake statistics, Smith proposes, is seismic moment. As defined in the dislocation theory of faulting, the seismic moment of an event is the product of rigidity, fault area and the average slip over the fault during the event. This relationship was exploited some years ago by Brune (*J geophys. Res*, 73, 777; 1968) who used it to estimate the total slip along major fault zones by summing the seismic moments of recorded earthquakes. Smith now reverses the process by using the geological record of total fault slip to determine the total historic seismic moment on a fault This total moment is then fed into an expression for maximum magnitude obtained by integrating the moments of all earthquakes along the fault up to the maximum magnitude.

Inevitably, such a process involves assumptions. One is the taking of an empirical relationship of the form $\log Y = c + dM$ between moment (Y) and magnitude (M). Another is the validity of the original N–M relationship. It must also be assumed that the

seismicity is a stationary process on a time scale of thousands of years and that during those years the maximum earthquake has occurred at least once. Finally, implicit in the method is the assumption that the observed geological displacement along the fault is due to earthquakes rather than creep, for any significant component of creep would lead to an overestimation of the maximum expected magnitude.

The need for these assumptions and natural deficiencies in the basic geological data combine to ensure that the method is less than perfect. On the other hand, as Smith points out, it does have the merit over existing techniques of making greater use of geological information. In any event, application of the method predicts maximum magnitudes of 8.2–8.4 for the first order branches (for example, the Calaveras fault) of the San Andreas fault system, 6.3–6.5 for lesser branches (for example, San Simeon) and 6.3 for still lesser branches (for example, West Huasna). Whatever the validity of these figures (compare magnitude 8.3 for the 1906 San Francisco shock) they are certainly more precisely defined than those obtained from fault length-magnitude curves. For a 200 km fault, for example, the conventional technique would only give an upper limit in the looser range 7.3–8.5. □

Poly(ADP-ribose)

from Mark Smulson and Sydney Shall

The fourth International Workshop on poly(ADP-ribose) was held in Hamburg on August 2–4, 1976, was sponsored by the Deutsche For-schungsgemeinschaft and organised by Professor Helmuth Hilz.

THE enzyme poly(ADP-ribose) polymerase requires DNA and catalyses the successive transfer of ADP-ribose units from NAD to histones and other proteins associated with chromatin, the net result being negatively charged, short polymers covalently attached to nuclear proteins.

The workshop concentrated on three main themes; possible biological roles for this modification, definition of the nature and chemistry of the protein acceptors of the polymer, and the purification and characterisation of the enzyme itself. The most exciting advance in the area was experiments suggesting that poly(ADP-ribose) cross-links chromosomal proteins. P. R. Stone and W. R. Kidwell (National Cancer Institute, Bethesda) reported that after incubation of HeLa cell nuclei with NAD, a dimer of histone H1 can be isolated containing one link-

ing chain of 15 ADP-ribose residues. Further evidence for this structure was provided in the previous week in a lecture at the International Congress of Biochemistry in Hamburg by G. Dixon (Calgary Medical School, Canada): he suggested either an inter-or intramolecular covalent linkage by way of poly(ADP-ribose) between glu-tamic acid residues near the amino terminal and carboxyl terminal ends of trout sperm histone H1. Since the con-centration of poly(ADP-ribose) in in-tact HeLa cells reaches a maximum at the S–G2 phase boundary, Kidwell sug-gested that the crosslinking property of poly(ADP-ribose) for histone H1 (and perhaps non-histone protein acceptors as well) might function to link widely spaced H1 molecules along internucleo-some regions and hence condense chromatin. This model was supported by the finding (M. Smulson, George-town University, Washington) that enzyme activity can be detected in internucleosome regions of chromatin and not on isolated nucleosomes.

Extensive purification of the enzyme has been accomplished independently in at least four laboratories. Okayama, Ueda and Hayaishi (Kyoto) have found during their 7,000-fold purifica-tion that an endogenous acceptor for ADP-ribose copurifies with the enzyme; and it may not be a simple protein. Histones decrease the K_m for NAD in their preparation. Confirmatory results were reported by K. Yoshihara (Nara Medical School, Japan). Both S. Shall (University of Sussex) and P. Mandel (Centre de Neurochimie du CNRS, Strasbourg) suggest the loss of an in-hibitor of the enzyme during their purifications. In addition, Shall and coworkers have estimated the K_i for a number of poly(ADP-ribose) poly-merase inhibitors which turn out to be also cyclic AMP phosphodiestrase in-hibitors, with higher affinity for the poly(ADP-ribose) polymerase, suggest-ing that some biological effects attri-buted to cyclic AMP from inhibitor studies, might really be due to poly-(ADP-ribose). Certain NAD analogues were reported to be incorporated into poly(ADP-ribose) chains (R. Suha-dolnik, Temple University).

Various new data on chromosomal protein acceptors for poly(ADP-ribose) were presented. M. G. Ord and L. A. Stocken (University of Oxford) find that histone H3 as well as H1 has ADP-ribose. Hilz and coworkers (Hamburg, Germany) have found heterogeneity both in the acceptors among non-histone proteins and in the sensitivity of the covalent linkage to cleavage. The proportion of monomer and poly-mer varies with changes in cellular proliferation rates. M. Miwa and T. Sugimura (National Cancer Centre, Tokyo) have been able to detect poly-

mer chains with equal (ADP-ribose) units but with different phosphate termini by acrylamide gel electro-phoresis.

One frustration has been the hetero-geneity in the "tentatively" reported covalent linkage of the first ADP-ribose to histones. The proposed link-ages include attachments from C-1 ribose to glutamyl carboxyl and threonyl hydroxyl by way of glycosidic bonds, to seryl phosphate by way of an ester bond or N-glycoside to arginine, and a possible Schiff base linkage. Much work clearly needs to be done in this area, but the heterogeneity of acceptors between histones and non-histone proteins suggests that vari-ous covalent attachment sites will ulti-mately be established.

The development of new ultrasensi-tive assays, including fluorometric determination (Mandel) and radio-immunoassays (Miwa and Sugimura; Bredehorst and Hilz; Kidwell) for quan-titation of poly(ADP-ribose) in cellular extracts, have produced substantial, but still far from definitive progress towards assigning a biological function to this fascinating chromosomal protein modification system. Reports at the workshop suggest that ADP-ribosyla-tion, like other nuclear protein modi-fications such as phosphorylation, probably has multiple functions all revolving around specific large or small structural pertubations of chromatin. Increases in enzyme activity are noted in SV40 transformation of cells (M. Miwa and T. Sugimura) for example; in induction of globin mRNA in Friend cells (E. Rastl, Ernst-Boehringer-Institut, Vienna) and an interesting correlation was described by A. Caplan (Case Western Reserve University, Cleveland) between intracellular NAD levels, poly(ADP-ribose) polymerase activity and the differentiation of mesenchymal cells either to muscle or cartilage.

Experimental approaches which make use of the effect of cytotoxic DNA-alkylating agents on poly(ADP-ribose) polymerase activity and *vice versa* also indicate that poly(ADP-ribose)–induced changes in chromatin structure may be important in repair of DNA damage. Fragmented sites seem to be a signal for increased *in vivo* activity of poly(ADP-ribose) poly-merase (groups at Georgetown and Sussex).

On the basis of data showing accu-mulation of poly(ADP-ribose) in mid-S phase and again in the G2 phase of the cell cycle (Kidwell) there was also speculation that chromatin repair and housekeeping must occur after the first peak of S phase, perhaps by way of polyADP-ribosylation relaxation of chromatin, and again in early G2 phase for cells to complete a normal cycle.

review article

Background of modern comet theory
Fred L. Whipple*

The modern understanding of comets dates from the publication in the early 1950s of three crucial ideas: that of a comet cloud around the Sun, that of the dirty-ice comet nucleus, and that of the "solar wind" to explain the comet tail. Starting from those three ideas, the observations of the past 25 years have led to the explanation of all the mysterious and sometimes spectacular phenomena associated with comets.

ANCIENT records indicate that extremely bright comets terrified almost everybody. Their erratic, unpredictable and mysterious activities seem to have touched off a basic human fear of the unknown, not just a fear of dangerous atmospheric objects which comets were assumed to be. This primitive fear, no doubt exploited by man's spiritual advisors throughout history, was still being exploited in this century. Entrepreneurs, for example, sold "comet pills" as preventative medicine against the noxious vapours from Halley's comet during its apparition in 1910. Chinese astronomers, incidentally, recorded Halley's great comet in 239 BC and probably as early as 467 BC[1].

Tycho Brahe of Denmark was the first to undermine the basis for this fear. He removed comets from the upper atmosphere and placed them in the sphere of modern scientific research by proving that the bright comet of 1577 moved in interplanetary space at a greater distance from Earth than the Moon. Sir Isaac Newton then applied gravitational theory to comets and calculated the first periodic orbit for a comet—that of 1680. Applying Newton's methods, Edmund Halley first predicted correctly the return of a comet, alas posthumously, when he identified the bright comet of 1682 as the return of comets seen in 1607 and 1531. His words: "therefore, if it should return according to our predictions about the year 1758 impartial posterity will not refuse to acknowledge that this was discovered by an Englishman." Hereby acknowledged!

Since about 1700, the discovery, the meticulous observing and the calculating of orbits for comets has been a major astronomical enterprise. Marsden's[2] *Catalogue of Cometary Orbits* lists 21 orbits in the seventeenth century; 63 in the eighteenth; 308 in the nineteenth; and 513 in the twentieth (until 1975). In all, the catalogue includes 964 apparitions of 625 individual comets of which 102 have periods less than ~150 yr. One must admire and be grateful for the patience, persistence and careful workmanship of the cometary astronomers who have given us so much invaluable basic information. But a knowledge of the precise orbits and appearance of many comets was only a necessary prelude to understanding their true nature. The phenomena of comets are produced by physical and chemical processes, which had to be ferreted out step by step. Some of these major steps are described in this paper.

Kepler, although he did not understand the orbital motion of comets, sought to explain earlier observations that comet tails point away from the Sun "by the supposition that a

tail is formed by rays of the Sun which penetrate the body of the comet and carry away with them some portion of its substance."[3]. He thereby anticipated the phenomenon of solar radiation pressure as a major factor in the formation of comet tails, although the phenomenon itself was not demonstrated theoretically until much later by Clark Maxwell, and in the laboratory by Lebedev in 1901. Bessel, in 1836, initiated the mathematical analysis of a solar repulsive force on particles ejected from comets and forced into their tails, although he looked on the force as electrical or magnetic. Bredikhin, in 1903, further developed and applied the theory to account for the observed highly curved dust tails. Theoreticians, however, remained frustrated in their attempts to explain motions in the great straight tails by means of solar radiation pressure alone. Accelerations exceeding a hundred times solar gravity, even on light ionised atoms or molecules, require a force that is much better coupled than solar radiation.

Schiaparelli, in 1866, made a major step in understanding comets when he identified the orbit of the Perseid meteor stream of August with that of Tuttle's comet of 1862, thus proving that comets lose solid bodies and therefore must contain them intrinsically. Two years earlier Donati made the first spectroscopic observations, a start towards the later identification of the bands and lines of radicals, atoms and ions and of the continuous spectrum reflected by dust. By 1911 Schwarzschild and Kron had identified the fundamental fluorescent process by which the band spectra are produced. Solar radiation is absorbed in ground energy levels of atoms or molecules, producing excitation. Fluorescence or reradiation then follows by cascading from high energy levels through various lower levels. Because of low space densities and dilute solar radiation the higher energy levels are erratically filled, depending on the intensity of the solar radiation at the specific wavelengths of transitions from the ground levels. Thus the relative intensities of lines in the band spectra are distorted from their laboratory counterparts. Swings'[4] demonstrated in 1941 that the variable Doppler effect of a comet's velocity with respect to the Sun changes the critical wavelengths absorbed from the solar spectrum by cometary particles. Irregularities in the spectra then produce variations in the cometary spectra so that the distorted bands change their appearance with time.

Thus in the 1940s it was recognised that bright comets such as Halley's near perihelion (0.59 AU: 1 AU=the Earth–Sun distance) lose the order of $10,000\,t\,d^{-1}$ of C_2 alone[5]. The total loss rate had to be increased by the other observed contributing radicals, such as CN, CH, CO, OH and NH, amounting altogether to a moderate fraction of a

* Address: Center for Astrophysics, Harvard College Observatory and Smithsonian Astrophysical Observatory, Cambridge, Massachusetts 02138.

ton per second! Estimates of the meteoroidal losses as evidenced by meteor streams were smaller but of comparable magnitude, while the spectra of meteors showed the generally abundant heavier elements Fe, Ca, Mg, Mn, Cr, Si, Ni, Al and Na. Meteorites, however, did not seem to be of cometary origin as no falls had been observed from cometary meteor streams.

Erratic variations in cometary brightnesses had always been known, and careful observations had shown that the relationship between intrinsic brightness and solar distance (r) varied greatly from comet to comet, the average relationship being $\sim r^{-3}$ to $\sim r^{-4}$. A few comets had disappeared while others had split into two or more distinct components accompanied by a "burst" of luminosity, usually to disappear later. For the Sun-grazing family, however, a major component usually persisted to a considerable distance from the Sun. Encke's comet had made more than fifty revolutions with a supposedly (but erroneously high) diminution of some ten times in brightness.

Most of these facts and phenomena fit qualitatively with Proctor's late nineteenth century concept of the comet nucleus as a sand bank or gravel bank. He pictured the fundamental particles as solid meteoroids from which absorbed gases are released by solar heating, the rate of gas desorption increasing rapidly with decreasing solar distance. The observed phenomena are evanescent, resulting from solar radiation acting in various ways on the gas and dust.

Attempts to define such a gravel bank quantitatively, however, led generally to increasing difficulties or contradictions. The nature of the dilemma can be seen qualitatively. One extreme is the nearly transparent gravel bank with negligible gravitational cohesion. The particles are all exposed to sunlight and must lose most of their gases on the first heating. Unless metres in dimension, they must be vaporised and entirely lost for the several Sun-grazing comets. Because each particle moves in its own orbit about the Sun, particles of different sizes will be forced into orbits of different periods by the Poynting–Robertson effect of solar radiation. Tus comets, especially periodic comets, should be elongated along their orbits if their nuclei are transparent gravel banks.

At the other extreme, the gravel-bank model is opaque near the centre and massive enough to hold particles gravitationally against solar tide-raising disruption. Because of orbital motions the particles will collide frequently at very low velocities and there is no reason to believe that the collisions will all be elastic. Thus a compact core will develop and become essentially asteroidal in character.

Intermediate gravel-bank models have been designed to meet some of the observational requirements only to stumble in explaining others. A possible hierarchy of such models to account for various types of comets is still no solution because even the fit of a special model to a specific comet is usually unsatisfactory.

Furthermore, a sinister persistent threat to the various gravel-bank models lay in Encke's (1823) discovery that the comet of shortest period was measurably decreasing in period, some 2 h in 3.3 yr. During the following century, this comet, which had been named after Encke, slowly reduced its rate of period decrease, that is, its deviation from gravitational motion. The popular postulate of a resisting medium between the planets, tempting as it may have been in the eighteenth century, became progressively less tenable in the nineteenth because no other supporting evidence could be found and because improved knowledge of interplanetary space impressed progressively lower limits on its possible density. The death knell for the idea sounded when two other periodic comets showed positive accelerations in their non-gravitational motions. A resisting medium can only reduce the dimensions of an orbit and thus the period. Increasing periods negate the possibility of a resisting medium. Conclusion: some comets must engender intrinsic

forces to change their periods, an idea suggested by Bessel[6] as early as 1836 and by Dubiago[7] in 1948, but without realistic physical mechanisms.

In spite of increasing detailed knowledge about comets and the abject failure of the gravel-bank model to fit the data quantitatively, the quantum jump in cometary theory was delayed until the middle of this century. Three explicit new ideas then revolutionised our concepts of comets The application and development of these ideas led to remarkable advances. These new theories are presented here in the order of their publication and are identified in terms of the cometary phenomena involved.

Oort's great comet cloud

In 1950 Oort[8] developed the theory of a great comet reservoir, or cloud about the Sun, providing a source of "new" comets to replenish the observable supply. Oort demonstrated that passing stars can perturb comets from huge orbits, extending out perhaps to 40–50,000 AU, into orbits with perihelion passages among the planets. Perturbations by the planets, particularly Jupiter, then eliminate more than half of such comets in successive passages, but disturb some into the short period orbits associated mostly with Jupiter. Oort thus defined specifically "The Home of the Comet"—a vague concept predating him by more than half a century. Although Opik[9] had long since shown that such a solar attached cloud is quasi-stable with respect to stellar perturbations, he had not specified the retrieval process that brings "new" comets on nearly parabolic orbits into the inner Solar System for the first time.

The dirty-ice comet nucleus

Whipple's[10] theory that the nucleus of a comet is a compact mass of ices with intermingling earthy materials formed a basis for understanding many aspects of comets, including (1) comet activity by way of the sublimation of ices by solar radiation, (2) loss of large masses of gas and solids from long-lived comets, (3) observed deviations of comets from Newtonian motions, (4) the origin of comets by the aggregation of cosmically-abundant low temperature condensates from typical interstellar or stellar gases.

The solar wind and ion tails

Biermann's[11] concept—that solar corpuscular radiation, now generally known as the solar wind, could produce the structural forms and high accelerations observed in the ion tails of comets—provided the key that opened the door to understanding the one most conspicuous phenomenon of comets. These long, nearly straight tails sometimes extend for more than 10^4 km and show accelerations up to 100 times or more of solar gravity As we have noted, the explanation—solar radiation—that previously had served semiquantitatively to account for dust tails, curving strongly backwards in a comet's orbital plane, proved quite inadequate when applied to ion tails.

These three new concepts were accepted relatively soon by most of the astronomical community because they promised to fill conspicuous lacunae in the theory of comets. All the major questions regarding the nature and phenomenology of comets could at last be approached by observers, theoreticians and celestial mechanicians on the basis of reasonable working hypotheses, to be subjected to the scientific rigour of checking and improvement. The full blossoming of cometary astronomy then emerged. The subsequent outstanding developments of space science, coupled with highly improved optical, infrared and radio sensors, added a wealth of new information that could be quantitatively integrated with older data.

Here I outline the development of my own ideas about the nature of comets. I first note my faith in a basic general principle to be applied in exploring the frontiers of any field in science. One is attempting to find the way through a

swampy morass of mutually conflicting observations, interpretations and theories. The principle: seize on a few "indisputable facts" that appear most relevant; momentarily ignore the remaining data. On these solid "facts" build conceptual structures or working hypotheses that are self-consistent, strongly rooted in the best evidence available. Then attempt to expand each working hypothesis by fitting in the successively weaker material. Rejection of material must be rigorously monitored. Is it irrelevant? Is it erroneous? Is it misinterpreted because of incomplete theory or because of an unnecessary or erroneous assumption, either tacit or explicit? If none of these, another working hypothesis should be investigated.

The general knowledge of comets available to me in the 1940s is summarised above. But I was highly conscious of three other "solid facts" that had been published but may not have been known or considered important by others who were thinking about comets. The first fact was the clear cut evidence I[12] had amassed from photographic meteor studies that Encke's comet had made not just about fifty revolutions but actually thousands of revolutions. The second fact was the evidence from the polarisation of the Zodiacal Light that interplanetary space at the Earth's solar distance contains less than 10^3 electrons cm^{-3} and, therefore, even fewer atomic nuclei[13]. The third fact was Opik's[9] demonstration that a cometary cloud could be quasi-stable about the Sun to a distance of several times 10^4 AU. The third fact left me with little concern about the replenishment of the shorter lived comets, because Opik's work, it seemed to me, had implied this likelihood.

The critical factor to me was the large quantity of matter that comets lose to space. If any comet can persist for a thousand or more revolutions, coming to 0.3 AU of the Sun at perihelion, it must contain or acquire a substantial quantity of gas and solids (cubic kilometres?). Conceivably a gravel-bank comet might persist if there were a gas source in space to refuel the solids by absorption. A back-of-the-envelope calculation suffices to show, however, that the known gas content of interplanetary space is woefully inadequate. Given a few km^2 effective area of the gravel bank, based on the brightness of comets at great solar distances, and even a 1,000-km s^{-1} radial velocity for the solar-emitted gas, the "refuelling" situation is hopeless.

Other implications of the gravel-bank theory were totally unsatisfactory to me. As mentioned above, a heirarchy of particle sizes would conspicuously spread the gravel-bank along its orbit in a thousand revolutions. Also I could never visualise a satisfactory process to split a gravel bank in the manner required to account for comet splitting, nor, indeed, a mechanism to produce comet bursts.

A dirty-ice model of the nucleus could, however, carry any specified reserve of solids and potential gas. It could lie dormant at great solar distances and become increasingly active as it approached the Sun. It could withstand the temporary heating experienced by the Sun-grazing comets and, indeed, split tidally to give birth to the family. It could account, at least qualitatively, for almost all of the many phenomena of comets. Furthermore, is was reasonable in terms of cometary origin as a low temperature condensate from typical solar or interstellar material. The most abundant compound-forming elements in a "solar mix", H, C, N and O provide the atoms in the radicals producing cometary fluorescence, while the heavier elements would certainly freeze out if the more volatile gases could. In principle, then, the basic difference between the conditions for the origin of the terrestrial planets and the comets become one of temperature.

At this stage of my thinking I felt that the icy conglomerate model was so obvious that it must already have been tacitly accepted by the major students of comets. There seemed to be little justification for publishing the idea unless some quantitative aspects could be developed. In a remark-

able treatise Hirn[14] had essentially demolished the idea of a resisting medium because of its resultant effect on comets' tails, and concluded that cometary gas would be frozen at aphelion to *"forme alors un assemblage de corps solides désagrégés."* Pol Swings, indeed, who was impressed by my evidence for the great age of Encke's comet, almost presented the icy model in 1948 (ref. 15). He concluded that the gas of a comet occurs by evaporation, not by desorption, but failed to develop the concept further, being constrained by the classical assumption that a comet nucleus consists of numerous small bodies and that breakage is important to produce activity. In retrospect it seems that this persistent assumption of a multibodied comet nucleus was the major obstacle to an earlier development of the icy concept. I was not so constrained because I realised that solids imbedded in a small icy body of low surface gravity would be blown out aerodynamically by sublimating gases. The novelty of this idea did not occur to me at the time and thus I abandoned the comet problem for a year or two.

I decided to publish the icy comet model while working on a theory (unpublished) of meteor drag in the atmosphere. I was attempting to include a term in the classical drag theory to allow for the jet action of vaporising meteoroidal material. Suddenly I realised that an icy comet nucleus should produce just such a jet action. Furthermore, any delay in the sublimation process for a rotating nucleus would produce a jet force component normal to the solar direction and thus accelerate or decelerate the orbital motion, depending on the sense of rotation. Here was a quantitative mechanism for explaining non-gravitational motions in comets with either positive or negative rates of period change. Rotation might be expected for any solid body in space, but certainly could be produced in an irregularly shaped icy nucleus. The integrated force vector of the jet action need not point precisely at the centre of gravity. Thus the model could explain comet splitting as a spinup phenomenon, as well as a tidal disruption.

A major problem of comets was and, to some extent still remains: the identity of stable "parent" molecules to be dissociated into the observed radicals. Not being a chemist I was content with the previously suggested H_2O, CH_4 methane- and NH_3 (ammonia) to account for OH, CH and NH, respectively, with CO_2 (carbon dioxide) or CO for CO. The radical CN was more difficult because the obvious C_2N_2 (cyanogen) is so active. I favoured HCN, or C_2H and N_2 requiring some chemical action to produce CN. The chemists were unhappy with these "non-equilibrium" chemical companions but I felt (hoped) the compounds might bed together at very low temperatures. The extremely low temperature (≤ 20 K), required to freeze out CH_4, CO or N_2 worried me more. This problem was resolved by Delsemme and Swings[16] in 1952 who suggested that the highly volatile CH_4 and others might be imbedded in H_2O snow as hydrates or clathrates. Their suggestion fits the observations well and has generally been accepted.

The most likely parent molecule should be H_2O because of its high heat of vaporisation. It should also be the most abundant. The luminosity variation of short period comets clearly indicated its presence. At a solar distance of 2.5–3 AU the vapour pressure of H_2O in sunlight becomes negligible as does the activity of most of these comets.

Calculations[10] based on the assumption of icy conglomerate nuclei with the compositions suggested above led immediately to reasonable jet forces and lags in emission for Encke's comet and others, requiring nuclei with radii of only a few km. The loss of matter needed to account for the non-gravitational forces was only a fraction of one per cent per revolution. For longer period comets the jet force radial towards the Sun should exceed the transverse lag component. This has since been demonstrated[17] as well as non-gravitational motions among most of the periodic comets[18].
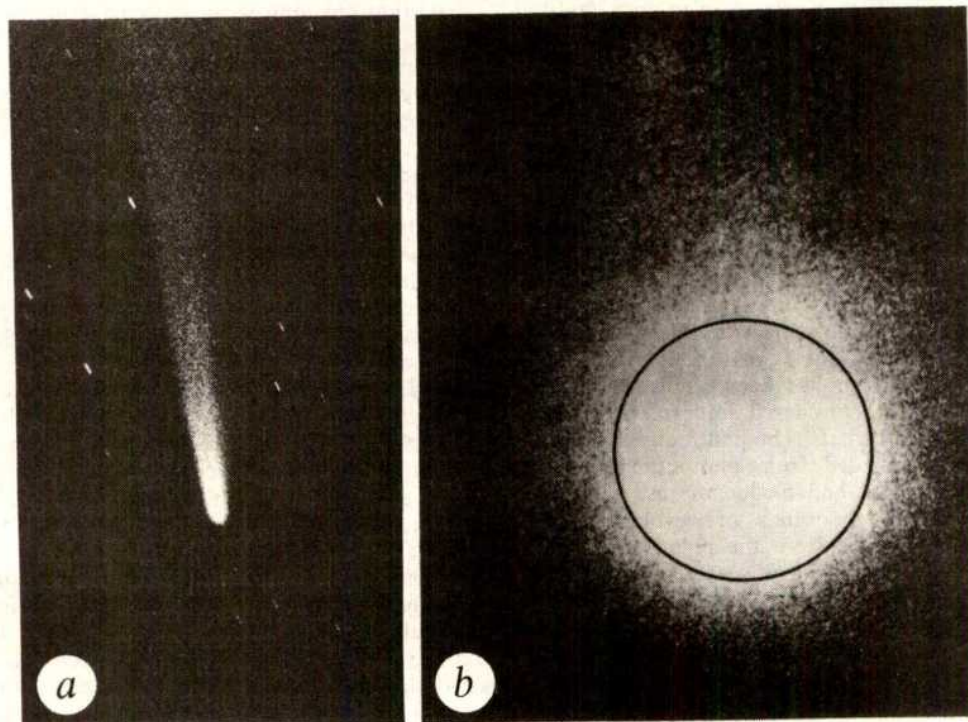
Fig. 1 Comet Kohoutek, 1973XII, Dec. 25.9, 1973. *a*, Photograph, Comet and Asteroid Observatory, South Baldy, New Mexico. *b*, NASA Skylab photograph in far-ultraviolet hydrogen Lα radiation. Both photographs are on the same scale as is the Sun's disk (circle) at same projected distance.

The great loss of H_2O from comets to be expected from the icy model was explored by Biermann and Trefftz[19] in 1960, who predicted a resultant huge hydrogen cloud observable by its Lyman α-emission line in the far ultraviolet (at $\lambda = 1,416$ Å). The observations from space vehicles dramatically verified this prediction for Comet Bennett, 1970II (refs 19, 20). The hydrogen cloud about Comet Kohoutek, 1973XII, in Fig. 1 dwarfs our old ideas about the spatial extent of bright comets. The total loss of H_2O amounted to a significant fraction of a 1 km³, typical of brighter comets.

Even though Comet Kohoutek may have been a dud in the eyes of the public, it was a huge success scientifically because it was the best observed comet in all history. Radio observations in the millimetre range produced spectacular new results : the identification of hydrogen cyanide, HCN (ref. 22) and methyl cyanide, CH_3CN (ref. 23). Methane and ammonia still remained elusive but the H_2O positive ion was identified in the optical spectrum and atomic O and C in the far ultraviolet. The neutral H_2O molecule resisted discovery until Comet Bradfield, 1974III. Comet Kohoutek brought the theorists to recognise the complex possibilities of gas-phase chemistry in the head of a comet, which leaves uncertain the true identity of some parent molecules because some of the radicals may be split up or chemically changed in the cometary atmosphere.

The presence of chemically ill-mated carbon molecules and the apparent low abundance of CH_4 in Comet Kohoutek show clearly that comets are a mixture of low temperature compounds, not formed by slow cooling from a hot gas. Such carbon compounds are prevalent in interstellar gas so that their occurrence in comets suggests a similarity between comets and the interstellar medium. Most students of the subject believe that comets probably formed in the outer parts of the planetary system, but do not exclude direct contributions from the interstellar cloud that collapsed to form the Solar System, or even the bare possibility of comet accretion in fragmented interstellar clouds connected gravitationally with the Sun.

The brilliant dusty Comet West, 1975n, split into at least four components, giving us an intimate look at a discrete comet nucleus. Figure 2 presents five successive large scale photographs of the head made at the New Mexico State University with their 20-inch reflector. The components are behaving as would be expected for multibodied nuclei: they are separating! The relative velocities are small, however, only a few m s⁻¹. But more interesting are the brightness variations shown on these beautiful photographs. At times components D, B and C are as bright or brighter than componet A. From an analysis of the motions, Sekanina[24] finds that component A is the main body of the comet moving in nearly a gravitational orbit. Component D separated first on February 13 starting at zero relative velocity with a relative deceleration, $d = 2.85 \pm 0.03$ in units of 10^{-5} solar acceleration. Component B separated second with $d \sim 5$ and C third with $d \sim 40$, an order of magnitude greater deceleration than D and B. As Sekanina predicted, C was short lived, becoming unobservable after three weeks from its separation on March 5.

Note the remarkable paradox. Component C must surely have been very much smaller and less massive than component A because it disappeared rapidly and because it displayed a huge non-gravitational deceleration, indicating a very much larger area–mass ratio for jet action. Nevertheless, component C outshone component A on at least two occasions. An explanation of this phenomenon lies inherent in the icy comet model. Let us accept the concept that the interior of an active comet contains a significant fraction of materials much more volatile than water ice. The surface layers, vaporising in sunlight, must become stratified with non-volatile dirt on the surface, water ice below, and successively more volatile materials underneath. The rate of mass loss is tempered by the dirt and hard-to-vaporise water ice. An exposed inner surface, however, will become far more active because of the materials with low heats of vaporisation.

The key to our explanation now lies in the simplistic fact that a solid of any shape whatsoever, when broken in

March 24.5

March 18.5

March 14.5

March 12.5

March 8.5

1976

Fig. 2 Nucleus of Comet West, 1975n, on 5 dates. Component A, lower left; B, upper right; C, lower right; D, middle left. Scale on March 24.5: A to B = 16.3″. Distance from Earth: 0.88 to 1.09 AU, March 5.5 to 24.5. From A. S. Murrell and C. F. Knuckles, New Mexico State University Observatory.

two pieces, exposes identical interior surface areas for each piece. Thus the small piece of a split comet nucleus exposes as much new surface containing highly volative materials as does the large piece. In certain orientations with respect to the Sun it can produce more activity than the large piece and hence outshine it. Continued breakage of the smaller component undoubtedly adds to its activity, to its rapid demise, and to its high decelerating force by jet action. Thus Comet West demonstrated before our eyes one of the striking properties of a dirty-ice comet nucleus.

As to the cause of comet splitting, other than tidal disruption, we have no definite proof. Rapid rotation produced by asymmetric jet action with resultant spinup seems to provide a reasonable explanation. It is not yet certain whether the equatorial layers need be spun off against gravity or whether the weak gravity, perhaps reduced by rapid rotation, allows the underlying volatiles or pockets of volatiles to eject large volumes of surface. In any case, splitting begets more splitting as would be expected if rotation has a role. A body in space will eventually rotate about the axis with maximum moment of inertia, being in unstable rotation about other axes. Loss of material usually changes the distribution of inertia so that after a split, the body will seek a new axis of rotation and alter the internal forces. More splitting may be expected and is frequently observed.

Plasma tails present a fascinating complex problem area for theoretical study and, indeed, have become a kind of space laboratory for magnetohydrodynamics. When instrumented space probes finally reach comets, the cometary space laboratory will become a reality, to answer many of these physical problems, as well as those of gas-phase chemistry in comets and, most important, to tell us about the true composition and physical structure of comet nuclei. From space probes we can hope to pin-point the circumstances of cometary origin and their relation to other bodies in the Solar System, particularly the Earth. I am convinced that comets contributed significantly to the life-giving volatiles of the Earth and, indeed, would not be surprised to learn that material contributed by comets made possible our very existence on this planet.

1 Ho Peng Yoke, *Vistas in Astronomy*, **5**, 127 (1962).
2 Marsden, B. J., *Catalogue of Cometary Orbits* (Smithsonian Astrophysical Observatory, Washington, 1975).
3 Berry, A., *A Short History of Astronomy*, 193 (Scribners, New York, 1910).
4 Swings, P., *Lick Observatory Bulletin No. 19*, 131 (1941).
5 Wurm, K., *Mitt. Hamburg. Sternw., Bergedorf.*, **8**, 57 (1943).
6 Bessel, F. W., *Astr. Nachr.*, **13**, 345 (1836).
7 Dubiago, A. D., *Russ. Astron. J.*, **25**, 361 (1948).
8 Oort, J., *Bull. Astr. Insts. Netherlds.*, **11**, 91 (1950).
9 Öpik, E. J., *Proc. Am. Acad. Arts Sci.*, **67**, 169 (1932).
10 Whipple, F. L., *Astrophys. J.*, **111**, 375 (1950); *ibid.*, **113**, 464 (1951); *ibid.*, **121**, 750 (1955).
11 Biermann, L., *Z. Astrophys.*, **29**, 274 (1951).
12 Whipple, F. L., *Proc. Am. Acad. Arts Sci.*, **83**, 84 (1940).
13 Whipple, F. L., and Gossner, J. L., *Astrophys. J.*, **109**, 545 (1949).
14 Hirn, G. A., *Constitution de l'espace Céleste*, 254 (Gauthier-Villars, Paris, 1889).
15 Swings, P., *Ann. Astrophys.*, **11**, 124 (1948).
16 Delsemme, A. H., and Swings, P., *Ann. Astrophys.*, **15**, 1 (1952).
17 Hamid, S. E., and Whipple, F. L., *Astrophys. J.*, **58**, 819 (1953).
18 Marsden, B. J., Sekanina, Z., and Yeomans, D. K., *Astr. J.*, **78**, 211 (1973).
19 Biermann, L., and Trefftz, E., *Z. Astrophys.*, **49**, 111 (1960).
20 Code, A. D., Houck, T. E., and Lillie, C. F., *IAU Circ.*, No. 2201 (1970).
21 Bertaux, J. L., and Blamont, J. C., *C. r. hebd. Séanc Acad. Sci., Paris*, **B270**, 1581 (1970).
22 Huebner, W. F., Snyder, L. E., and Buhl, D., *Icarus*, **23**, 580 (1974).
23 Ulich, B. L., and Conklin, E. K., *Nature*, **248**, 121 (1974).
24 Sekanina, Z., *IAU Circ.*, No. 2943 (1976).

articles

Transitions in double-diffusive convection

Herbert E. Huppert

Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge CB3 9EW, UK

The transitions of solutions of differential equations from one stability regime to another are of great current interest, both mathematical and physical, and there are conflicting hypotheses as to how such transitions occur. Here we present the results of an investigation of double-diffusive convection, which is important in oceanography, astrophysics and chemical engineering. The calculated transitions are found to be very different from those previously suggested.

MOST fluid motion is turbulent. One approach that has been used in investigating this form of motion is to examine the transitions undergone by an initially laminar flow in its evolution to a turbulent state. Landau[1] hypothesised that as some appropriate non-dimensional parameter, R, increases, a critical value is reached at which the particular time-independent laminar flow under investigation becomes unstable and equilibrates to a new time-dependent flow. At a larger critical value this flow itself becomes unstable and this process of transitions is envisaged to continue until the flow has become so complicated that it would be referred to as turbulent. An alternative hypothesis has been recently suggested by Ruelle and Takens[2], whereby after a relatively small number of transitions, a well defined value of R is reached at which the flow exhibits an abrupt transition to a far more complicated, random, and hence turbulent, motion. Since these are both abstract hypotheses, only a direct investigation of the governing equations of motion can decide if either of these descriptions is correct, or if a different set of transitions takes place. In an investigation of this sort, but motivated by biological considerations, May[3] recently determined the transitions that occur in a number of first-order nonlinear difference equations. He found that as the appropriate R increases through a sequence of critical values, the solution changes from consisting of one stable equilibrium point, through stable cycles of period 2^n ($n=1,2,\ldots$), and then, at a finite value of R, there is chaos, with slightly different initial conditions leading to solutions which diverge with time. We present here the results of explicit calculations of the transitions that occur in two-dimensional double-diffusive convection. The motion is governed by a set of coupled nonlinear partial differential equations and the transitions that occur are different in form from the three discussed above.

Double-diffusive convection

Double-diffusive convection is a generic term for the type of convection that occurs in fluids in which there are two components of different molecular diffusivities which contribute in an opposing sense to the vertical density gradient. For different sets of components, this form of convection has an important role in oceanography, astrophysics and chemical engineering[4]. Here we use the terminology of heat and salt, the components appropriate to oceanography. We restrict attention to the

Rayleigh–Bénard problem[4], where the fluid is considered to occupy the space between two infinite planes separated by a distance D, with the upper plane maintained at temperature T_0 and salinity S_0 and the lower plane maintained at temperature $T_0+\Delta T$ and salinity $S_0+\Delta S$ (ΔT, $\Delta S > 0$). We assume both planes to be stress free and perfect conductors of heat and salt, and restrict attention to two-dimensional motion, dependent only on one horizontal coordinate and the vertical coordinate. Expressing all lengths in units of D, time in units of D^2/κ_T (where κ_T is the thermal diffusivity) and representing the temperature T^* and salinity S^* by

$$T^* = T_0+\Delta T(1-z+T); \quad S^* = S_0+\Delta S(1-z+S)$$

we can write the governing Boussinesq equations of motion in terms of a stream function ψ as

$$\sigma^{-1}\nabla^2\partial_t\psi - \sigma^{-1}J(\psi,\nabla^2\psi) = -R\partial_x T + R_S\partial_x S + \nabla^4\psi$$
$$\partial_t T+\partial_x\psi - J(\psi,T) = \nabla^2 T$$
$$\partial_t S+\partial_x\psi - J(\psi,S) = \tau\nabla^2 S$$
$$\psi = \partial_{zz}^2\psi = T = S = 0 \quad (z = (0,1))$$

where

$$J(f,g) = \partial_x f\partial_z g - \partial_z f\partial_x g$$

Fig. 1 The results of linear stability analysis for $\sigma = 1$ and $\tau = 10^{-\frac{1}{2}}$. Overstability first occurs along $R = R_1$ and monotonic stability first occurs along $R = R_6$. Only the hatched region is stable to linear disturbances.
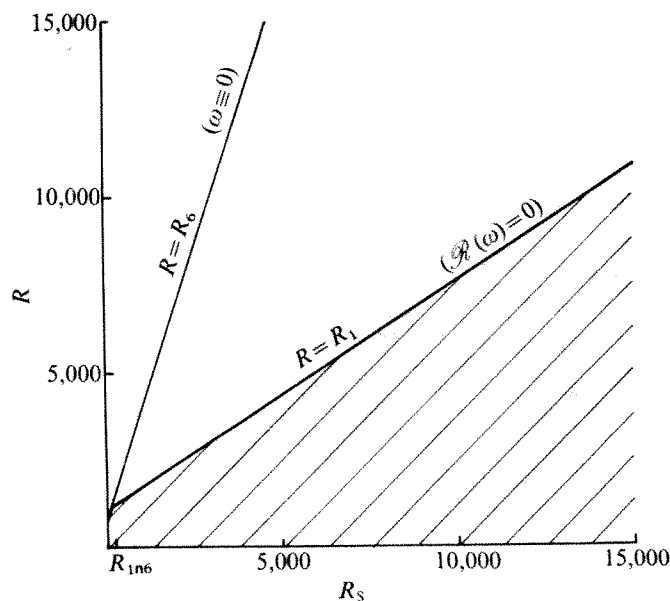
The four non-dimensional parameters appearing in these equations are: the Prandtl number $\sigma = \nu/\kappa_T$, where ν is the kinematic viscosity; the ratio of the diffusivities $\tau = \kappa_S/\kappa_T$, where κ_S, the saline diffusivity, is less than κ_T; the thermal Rayleigh number $R = \alpha g \Delta T D^3/\kappa_T \nu$, where is α the coefficient of thermal expansion, and g is the gravitational acceleration; and the saline Rayleigh number $R_S = \beta g \Delta S D^3/\kappa_T \nu$ where β is the saline analogue of α.

Stability

The criteria for the linear instability of these equations are obtained by neglecting the nonlinear Jacobian terms and representing the solutions in terms of the lowest normal modes with an exponential time dependence of the form $\exp(\omega t)$. The onset of overstability, defined by $\mathscr{R}(\omega) = 0$, first occurs for a value of R which we denote by R_1 and exchange of stabilities, defined by $\omega = 0$, first occurs for a value of R which we denote for convenience by R_6. In both cases the horizontal wavelength of the instability is $2^{3/2}$. In the $R-R_S$ plane the linear stability boundary is a combination of $R = R_1$ and $R = R_6$, as depicted in Fig. 1, which presents a summary of the linearised results for $\sigma = 1$ and $\tau = 10^{-1/2}$. For $R_S > R_{1n6}$, where R_{1n6} is the value of R_S at which $R_1 = R_6$, as R exceeds R_1 the conduction state $(\psi = T = S = 0)$ becomes unstable to an oscillatory mode. For $0 < R < R_{1n6}$ as R exceeds R_6 the conduction state becomes unstable to a monotonic mode.

We present here the form of the solutions of horizontal wavelength $2^{3/2}$ which emanate from the above linear transitions, or bifurcation points, as the degree of nonlinearity increases. An appropriate way to describe the resulting solutions is as a function of their amplitude. This is conveniently represented by either the thermal or saline Nusselt numbers evaluated by the lower boundary

$$N_T = 1 - \overline{\partial_z T}|_{z=0} \quad \text{and} \quad N_S = 1 - \overline{\partial_z S}|_{z=0}$$

where the overbar denotes a horizontal average, or by their respective temporal maxima, M_T and M_S. Since for most physical systems $R_S > R_{1n6}$ and this is conceptually the most interesting case, we concentrate on the latter regime and, principally by numerical integration of equations (1), map out the forms of equilibrium solutions in an $R-M_S$ plane in the manner depicted in Fig. 2. The discussion is in general terms; specific details and physical interpretations of the solutions will be published elsewhere (H. E. Huppert and D. R. Moore, in preparation).

The bifurcation point at R_1 can be either supercritical (R increases as M_T or M_S increases) or subcritical (R decreases as M_T or M_S increases). By the straightforward use of modified perturbation theory[5], a very lengthy relationship can be determined (H. E. Huppert and D. R. Moore, in preparation) which indicates, for fixed σ, τ and R_S, which of the two possibilities occurs. It is known from general theory[5] that solutions on branches emanating from subcritical bifurcations are unstable until the branch reaches a minimum value of R. Thereafter the branch continues with the amplitude of the associated time-dependent solutions increasing with increasing R. Supercritical branches are known to support stable solutions. A typical plot of N_T and N_S against time for a solution on a stable portion of the branch sufficiently close to R_1 is shown in Fig. 3a.

Increasing R

As R increases, this form of motion continues until R reaches a specific value, R_2 say. At $R = R_2$ the solution changes in form and develops a further structure as is indicated in the form of N_T or N_S as a function of time, as plotted in Fig. 3b. In both N_T and N_S there are four extrema, two maxima and two minima, per period, where the period is defined in the usual sense as the time between two identical states. In modern mathematical jargon, the solution for $R < R_2$ is on a sphere, while the solution

R389
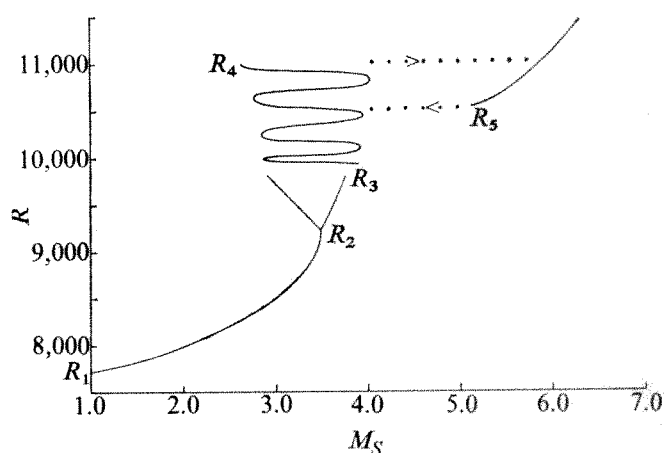
Fig. 2 The stable oscillatory branch and the stable monotonic branch for $\sigma = 1$, $\tau = 10^{-\frac{1}{2}}$ and $R_S = 10^4$. For $R_2 < R < R_3$ the branch is drawn so as to show both local maxima. For $R_3 < R < R_4$ the wavy line indicates that the solution is non-periodic and no definite maximum can be assigned. The dots at $R = R_4$ and $R = R_5$ indicate the jumps in M_S which occur as the solution changes from the oscillatory branch to the monotonic branch.

for $R > R_2$ is on a torus, and the transition at $R = R_2$ is called a bifurcating torus[6]. This form of motion continues until $R = R_3$ say, at which value a transition to a disordered solution occurs. A typical plot of N_T and N_S against time for such a disordered solution is shown in Fig. 3c. Long computation runs have not revealed any discernable periodic structure in the solution. Such non-periodic solutions continue to exist for increasing R until, for $R = R_4$ say, the only equilibrium solutions are time independent. For some values of σ, τ and R_S the transition at R_4 occurs before the one at either R_2 or R_3.

For all $R > R_4$ all equilibrium solutions are independent of time. Time-independent motions are more efficient at transporting heat and salt than time-dependent motions and thus the Nusselt numbers undergo a discontinuous increase as the solution changes from being on the oscillatory branch to the monotonic branch, as indicated in Fig. 2. As R increases

Fig. 3 The thermal and saline Nusselt numbers as a function of time for $\sigma = 1$, $\tau = 10^{-\frac{1}{2}}$ and $R_S = 10^4$. a, $R_1 < R = 8,600 < R_2$; b, $R_2 < R = 9,800 < R_3$; and c, $R_3 < R = 11,000$. The horizontal line in the bottom right-hand portion of c, represents non-dimensional time.
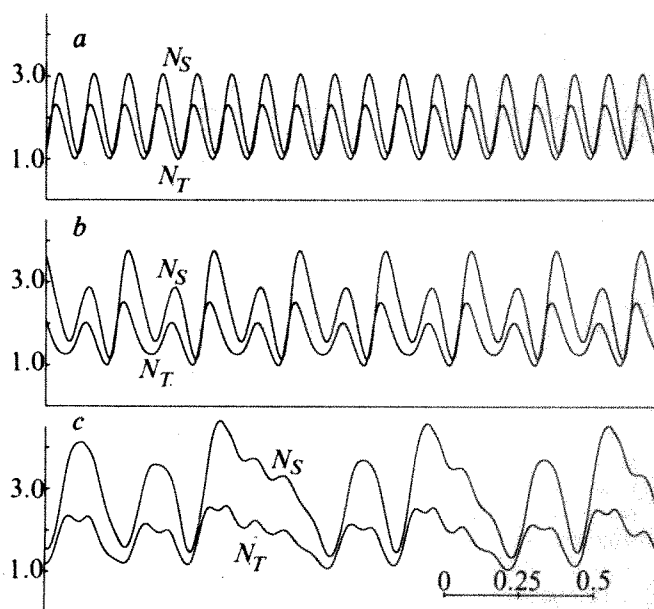
beyond R_4 the effects of the salt field decrease, and for $R \gg R_4$ the equilibrium solutions approach those for $R_S = 0$, a situation which has been intensively investigated by Moore and Weiss[7].

Decreasing R

If R is gradually decreased from some value greater than R_4, the equilibrium monotonic solutions retrace the states that would have been obtained on increasing R from R_4; for each $R > R_4$ there is a unique stable equilibrium solution. If R is decreased below R_4, an equilibrium time-independent solution continues to exist, with decreasing amplitude, until $R = R_5$ say. Further decrease of R leads to a solution on the oscillatory branch already described, or if $R_5 < R_1$ to the null solution. Thus, as indicated in Fig. 2 there is hysteresis between the two different forms of solution.

The time-independent branch of solutions emanates from the bifurcation point at $R = R_6$, which modified perturbation theory shows to be subcritical if $R_S > R_{1n6}$. As mentioned previously, solutions on such a subcritical branch are unstable until the minimum value of R is attained. Thereafter the branch continues and solutions on it are stable, with the amplitude of the solution increasing with increasing R.

If $R_S < R_{1n6}$ only time-independent equilibrium solutions are possible. For sufficiently small R_S the bifurcation at R_6 is supercritical, otherwise it is subcritical (H. E. Huppert and D. R. Moore, in preparation).

Conclusion

To summarise, in the most general case, as R increases there is a transition from the conduction state to an oscillatory motion (Fig. 3*a*), followed by a transition to a more complicated form of oscillatory motion (Fig. 3*b*), followed by a transition to a non-periodic, random state (Fig. 3*c*), followed by a transition to steady motion. Hence by increasing R it is possible in this situation to suppress disordered motions. For some values of σ, τ and R_S a sufficiently large disturbance can cause a transition directly from the conduction state to the steady state, while for other values of σ, τ and R_S only some of the intermediate transitions are omitted.

I thank Dr D. R. Moore, who wrote the numerical program used here, Dr Joyce Wheeler, who helped with the computations and preparation of the graphical output, and Professor R. M. May and Dr D. P. McKenzie for critical comments. This research was supported by the British Admiralty.

1 Landau, L. D., and Lifshitz, E. M., *Fluid Mechanics* (Pergamon, Oxford, 1959).
2 Ruelle, D., and Takens, F., *Commun. math. Phys.*, 20, 167–192 (1971).
3 May, R. M., *Science*, 186, 645–647 (1974); *Nature*, 261, 459–467 (1976).
4 Turner, J. S., *Buoyancy Effects in Fluids* (Cambridge, 1973); *A. Rev. Fluid Mech.*, 6, 37–56 (1974).
5 Sattinger, D. H., *Topics in Stability and Bifurcation Theory, Lecture Notes in Mathematics*, 309 (Springer, Berlin, 1973).
6 Hirsch, M. W., and Smale, S., *Differential Equations, Dynamical Systems, and Linear Algebra* (Academic, New York, 1974).
7 Moore, D. R., and Weiss, N. O., *J. Fluid Mech.*, 58, 289–312 (1973).

Stratified waters as a key to the past

Egon T. Degens

Geologisch-Paläontologisches Institut der Universität Hamburg, Bundesstrasse 55, 2000 Hamburg 13, FRG

Peter Stoffers

Laboratorium für Sedimentforschung, Mineralogisch-Petrographisches, Institut der Universität Heidelberg, Berliner Strasse 19, 6900 Heidelberg, FRG

Density stratification in lakes and oceans generate anoxic conditions below the pycnocline, and sediment facies mirror this development. A comparison of modern sediments deposited in stratified and non-stratified waters with sediments formed since the Cambrian reveals that the ancient sea has been stratified a number of times.

THE world ocean is fully oxygenated except for some local areas in regions of upwelling or stagnation. Oxygenation is related to modern climates, that is thermo- and haloclines, which can inhibit the mixing of water masses between stratified layers but which are only temporarily stable. Furthermore, cold polar surface water sinks and moves towards the equatorial abyssal plains. These and other climate-controlled physical factors turn over the ocean's water in a matter of a few hundred to a few thousand years. At this rate, molecular oxygen is recharged much faster in the deep sea than it is consumed by the oxidation of organic matter at greater water depths.

During prolonged warmer climatic stages or when land-locked seas develop, thermo- or haloclines may become so stabilised that they only move up and down in response to seasonal changes, tectonic activities, or some other major perturbations in the environment; but they rarely break up entirely. In such conditions molecular oxygen will remain abundant in the euphotic zone but will gradually drop to zero below the density boundary. This will cause the development of a euxinic environment in which no higher forms of life can exist and molecular oxygen is replaced by hydrogen sulphide.

It is important to know what happens to the strata when this situation arises because such information is crucial to the task of explaining the origin of euxinic sediments, which are so plentiful in the stratigraphic record. Unfortunately, comparative studies between shallow and deep-water habitats are hindered because most of the marine sediments exposed on continents are of shallow-water origin and suites of abyssal sediments have only recently become available through the Deep Sea Drilling Project.

A significant aspect of this problem which has received little attention in the past concerns a possible feedback mechanism between oxidising and reducing environments which may result in the formation of specific sediment types.

We will focus attention here on the feedback question by examining sediments of the same ages and geological settings which differ only in that one group has formed below and the other above a well defined thermo-halocline. The sediments are from cores from the Black Sea and some deep East African rift lakes and they represent continuous sections through parts of the Holocene and Pleistocene. A detailed examination of a thermo-halocline, oscillating through time, will allow us to follow in slow motion the impact of the reducing on the oxidising environment and vice versa.

Phase boundaries

An heuristic theorem implies that physicochemical phenomena established at the boundary between two different states characterise these two states. In the present context the theorem implies that the physicochemical properties of the reducing and oxidising environments are 'written' on the boundary layer, that is, the thermo-halocline. We will now examine this interface.

Mechanism and rate of molecular exchange across a well
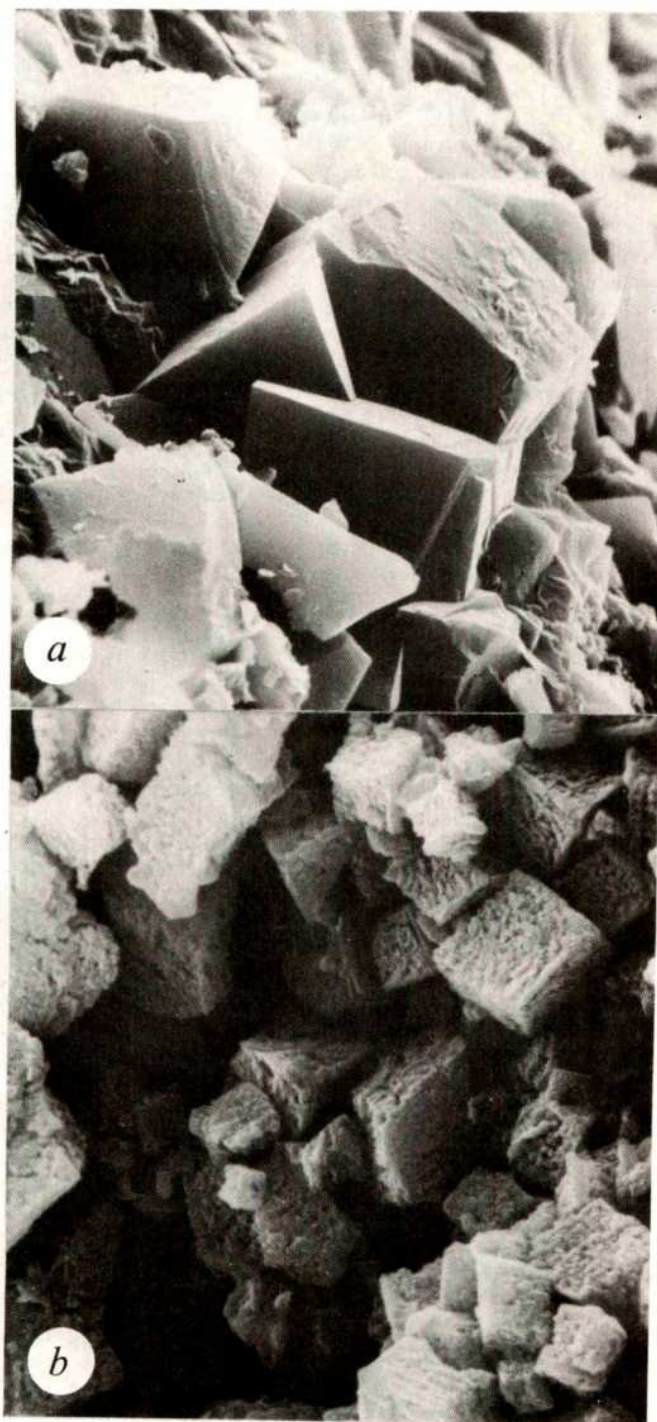
Fig. 1 Scanning electron micrographs of chemically precipitated calcites, sedimented above O_2–H_2S interface (*a*), and below interface (*b*). Size of individual crystals ~ 5–10 μm.

developed thermo-halocline have been studied thoroughly[1-3]. Vertical advection and diffusion will move deep water through the density boundary to the surface layer. For the main pycnocline in the Black Sea this results in a vertical eddy-diffusion coefficient of ~ 0.014 cm² s^{-1} (refs 2 and 3). From this value one can readily calculate the net upward flux of dissolved species such as iron, manganese, or hydrogen sulphide. Against the upward advective gradient there is a downward diffusion gradient of oxygen. These two opposing fluxes will generate a wide range in redox potentials within a narrow segment of the water column. Some elements and minerals will respond to this Eh gradient by becoming reduced or oxidised, and precipitated or dissolved, respectively. In turn, minerals can form at one point, but by sinking through an Eh gradient they may

redissolve. Depending on the speed of this reaction and the general hydrographic setting, the process of precipitation and dissolution can continue again and again and lead to substantial concentrations of certain elements and minerals close to the pycnocline.

Models using a constant eddy-diffusion coefficient for the vertical transfer of passive properties in a continuously layered medium[4,5] are only applicable for periods of hours and days because density boundaries tend to rise and fall and many other physical perturbations do exist[6-9]. If the interface is lowered rapidly, an almost spasmodic mineralisation of the upper layer will take place resulting in massive precipitation of a series of minerals. In contrast, a rise of the interface will dissolve the same minerals. Figure 1 shows calcite which has precipitated in the upper layer. In Fig. 1*a* the calcite remained in the surface layer, whereas in Fig. 1*b* it sank into the anoxic water. It is of note that all carbonates including dolomite and siderite will eventually dissolve if exposed to highly reducing environments.

East African rift lakes

The deep rift lakes of East Africa—Tanganyika, Malawi and Kivu—are thermally stratified. The perennial thermocline in Lake Tanganyika lies at ~ 100 m but is somewhat deeper and much more variable in Lake Malawi. Lake Kivu has a stable thermo-haline density structure with an interface at ~ 70 m (ref. 10). The position and nature of the density boundary influence the occurrence of distinct mineral and chemical facies observed in sediment cores of the three lakes[10-13]. Most notably affected by the position of the interface is the distribution of Fe-oxides, Fe-silicates, Fe-phosphates, Fe-sulphides, Mn-siderites and a series of Ca- and Ca–Mg-carbonates. In the following, we will examine sediments from deep (~ 500 m) and shallow (~ 50 m) parts of Lake Kivu that were deposited during the past 10,000–14,000 yr.

The impact of an oscillating stratification is illustrated by the distribution of manganese and related facies in a core from the central basin (see Fig. 2). The observed Mn^{2+} oscillations are interpreted as resulting from the following processes[14]. Thermal stratification allows Mn^{2+} and Fe^{2+} to accumulate in the water of relatively low pH below the thermocline. Vertical advection and diffusion transfer both ions through the interface where they instantaneously precipitate as manganosiderite. The newly formed mineral phase sinks back into deep water, dissolves, and the cycle starts over again. Its incorporation into the sediments will only be possible near the thermocline. The manganosiderite in Lake Kivu occurs, however, as discrete layers rather than as a mixture with the regular detritus. This suggests a more powerful transfer mechanism than advective eddy-diffusion: the rapid lowering of the thermocline and a resultant massive precipitation of the manganosiderite. Fluctuations in Mn^{2+} content shown in Fig. 2 reflect oscillations of the thermocline which shifted the area of manganosiderite deposition away from or back towards the location of the core site. High water stands are indicated by the emergence of sulphides and specific diatom assemblages in an organic-rich facies free of carbonates[14,15].

The profile continues in Fig. 3. This time, however, carbonate and organic carbon serve as indicators for the oscillating O_2–H_2S boundary[10]. Carbonate precipitation starts with high Mg-calcites and the percentage of $MgCO_3$ in solid solution is indicated numerically. At $CaCO_3$ concentrations > 10%, aragonite occurs. A change in mineralogy from Mg-calcite to aragonite reflects an increase in the Mg–Ca ratio in the lake water[13] which can best be explained by progressive evaporation which rapidly will lower both the interface and the water level and enforce lime extraction. The emergence of carbonate-free sapropels signals meromixis. Surface waters have pH values between 8 and 10 but as the pH drops below 7 the waters become anoxic. Carbonate particles falling into the deep water redissolve, whereas those that settle in shallow areas above the pycnocline remain intact and form thick sediment sections. Changes of the Mg–Ca ratio in the lake water will

Fig. 2 Distribution of manganese in Kivu sediments deposited during the time interval between 13,700 and 11,000 yr BP. Mineral facies and reconstructed thermocline oscillations are shown. High water stands are indicated by the emergence of sulphides and specific diatom assemblages[14,15].

established ecosystem. Whole populations can be exterminated; others can be pushed into restricted ecological niches from where they may perhaps reoccupy their former territory should the habitat change back to its former condition. Some species may adapt to the new environmental circumstances, whereas others may take advantage of the new situation and become dominant. In addition, some new species may evolve.

Lake Tanganyika, which has had a series of oscillating thermoclines and water levels[19-22], is a good test case for the evolution of species in a closed system. A large number of endemic genera and species is present in the modern lake[23], especially fish. When counted in 1960 (ref. 24), *Cichlidae* alone numbered no less than 133 species from 38 genera. Another point of interest is the striking resemblance of a great variety of present-day snails to certain marine forms that became extinct in the Jurassic[25,26]. All this has to be viewed against a background of a lake which formed in Miocene time and has never been open to the sea. In conclusion, the antiquity of Lake Tanganyika is mainly responsible for the evolutionary diversification, but temporal variability in lake levels, water chemistry, and stratification are considered dominant factors in triggering this development[14,15].

Black Sea

Much has been learned about the recent history of the Black Sea through a R/V Atlantis research cruise conducted by the Woods Hole Oceanographic Institution in spring 1969 (ref. 27). A topic of special significance for our work is the evolution of anoxic conditions during the Holocene[28].

Each gram-atom of carbon in plankton requires 1.3 mol of oxygen for its oxidation[29]. Since the amount of organic matter that falls through the O_2–H_2S interface is fairly well known[30], the 'decay' constant of the oxygen reservoir can be calculated[28]. From this number, the upward progression of the O_2–H_2S interface can be derived. Calculations show that the thermohaline boundary requires ~ 3,000 yr to rise from the 2,200-m-deep abyssal plain to a water depth of 500 m. Varve-counting techniques substantiate these findings[31] and indicate that finely laminated sapropels began to occur on the basin slope at a depth of 470 m almost 2,300 yr after they first appeared on the bottom of the Black Sea basin.

The sapropel facies is low in carbonates and many grains show corrosional features of the kind shown in Fig. 4. Of special note is the sapropel facies in a core from a water depth of 470 m which rests on a massive calcareous mud and is

determine the carbonate mineralogy[13,16]. Calcite, Mg-calcite, aragonite, protodolomite and monohydrocalcite generate cyclic patterns in the shallow basins of Lake Kivu, while at the same time a thin sapropel facies forms in the deep anoxic basin of the lake. Similar relationships exist in Lake Tanganyika[11,13].

The impact of an oscillating density boundary or a 'catastrophic' overturn is registered by the established fauna and flora. For instance in Lake Kivu the fish fauna is depauperate, chaoborids are absent and the molluscs, especially pelecypods, have suffered recent extinctions, judging from the abundance of shells of species which are no longer extant in the lake[15,17]. Particularly illuminating is the distribution pattern of diatoms over the past 14,000 yr (refs 14, 15 and 18). *Stephanodiscus astrea* declines almost linearly from ~ 90% of the total diatom population at ~ 14,000 yr BP to 0 at ~ 5,000 yr BP. Similarly, *Nitzschia fonticula* disappears at ~ 5,000 yr BP. Only one species, *Nitzschia spiculum*, crosses this boundary and is joined by newly emerging diatom species. These and related observations strongly suggest that changes in water chemistry introduced by a fluctuating interface can severely disturb the

Fig. 3 Distribution of $CaCO_3$ and organic carbon in Kivu sediments deposited during the past 11,000 yr (continuation of profile shown in Fig. 2). Numbers on carbonate curve represent the mole percentage of $MgCO_3$ present in high Mg-calcite; water depth: 473 m (ref. 10).
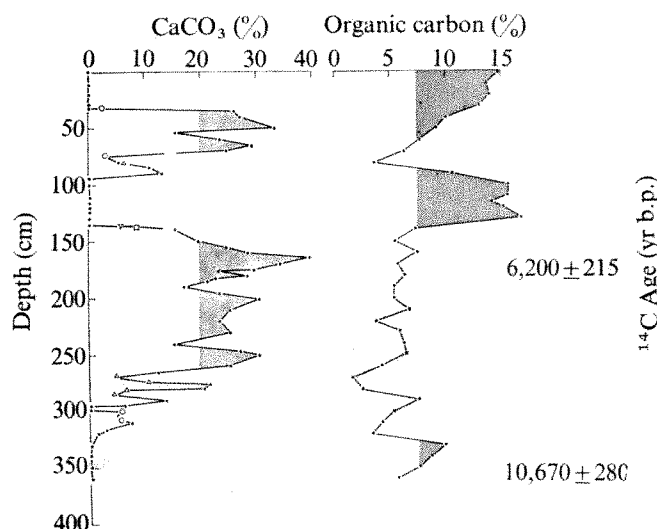
characterised by intercalations of very fine authigenic carbonate laminae up to 2 mm thick at the base of the sapropel but becoming finer and eventually disappearing towards the top of the sapropel layer. All these observations fit the general pattern described for Kivu sediments, and therefore indicate the same mechanism of formation.

Present—a key to the past

We have studied stratified waters in numerous environments for more than ten years[10–15,27,31,32]. It never occurred to us, however, that the results of these studies could significantly alter established concepts in geology. In other words, the subject was looked at more as an interesting anomaly rather than as a feature of global significance.

This outlook changed when we described sediment cores from the Deep Sea Drilling Project's Leg 42 B obtained by the Glomar Challenger in the Black Sea during May–June 1975. In three cores from the euxinic abyssal plain (core length 624.5 m) the basin apron (core length 1,073.5 m) and the basin slope (core length 503.5 m) a kind of textbook profile emerged consisting of successions of (1) shallow- and deep-water facies, (2) flysch and molasse cycles, (3) oxidising and reducing sediments, (4) clastic and chemical deposits, and (5) freshwater–brackish–marine series. Most striking is the distribution of authigenic carbonates. In going from the shallow to the deep basin, a typical facies succession is dolomite–aragonite–high magnesium calcite–calcite–siderite–carbonate–free sapropels. In view of these results, we will briefly compare present and past common sediment facies of the open marine environment. Since authigenic carbonates appear to be excellent environmental indicators, we will use them as tools in our comparison.

In modern marine environments, the chemical precipitation of calcium carbonates or dolomites is virtually nil. Recent marine carbonates are principally biological in origin. Calcitic foram and coccolith oozes cover wide parts of the deep sea and coral reefs, and coquina are found in some shallow-water environments. The distribution of calcareous oozes is controlled by the local carbonate compensation depth and a few other critical factors determining rates of dissolution[33]. In contrast, sand and clay are the typical facies in modern shallow-water deposits.

The stratigraphic record, from the Cambrian on, indicates an almost reversed situation. Hundreds of metres of shallow-water carbonates, distributed over extensive areas, are frequently encountered. Sedimentary cycles measuring not more than a few centimetres or metres can be traced over wide distances. On the other hand, an ordinary geosyncline commonly starts with a massive graywacke facies or a black shale facies free of carbonates. Judging by their physical appearance and chemistry, these sediments formed at great water depths and under anoxic conditions. In conclusion, sediment facies in the Black Sea and some East African rift lakes have many characteristics in common with ancient marine sediments, whereas sediment facies in the present ocean show little resemblance to the past record. These findings have interesting consequences because they suggest that the former ocean looked much like the present Black Sea. Two simple graphs will summarise the most critical points.

In the modern fully oxygenated ocean, sands and clays are deposited in the shallow sea (Fig. 5a). On stratification of the water, anoxic conditions develop below the interface and euxinic sediments form in the H_2S zone, whereas carbonates may develop in the oxic zone (Fig. 5b). It is essential to know that these two contrasting environments not only differ in oxygen and hydrogen sulphide but in a number of other chemical ingredients. As a common rule the anoxic waters are enriched in (1) mineral nutrients such as nitrate, phosphate, ammonia and silica, (2) some common elements, and (3) dissolved gases.

Lowering of the interface (Fig. 5c) will result in the transfer of dissolved chemical species into the upper layer. Primary productivity will be stimulated and carbonates will precipitate with the Mg–Ca ratio in the water phase controlling the carbonate
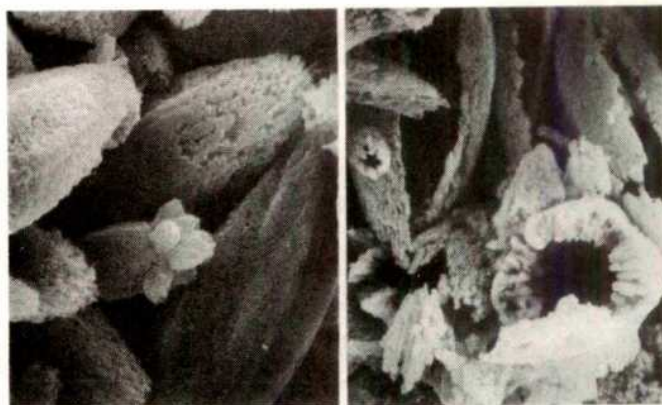
Fig. 4 Scanning electron micrographs of chemically precipitated aragonites buried below the O_2–H_2S interface. Note dissolution features similar to Fig. 1b. Size of aragonite grains ~ 20 µm.

mineralogy. An upward progression of the interface (Fig. 5d) will extend euxinic conditions to shallower parts of a basin, with established benthic communities becoming extinct, and carbonates starting to dissolve because the level of carbonate compensation in time will be identical with the level of the O_2–H_2S interface. As a function of climatic variation and constant reshaping of marine basins by tectonic or erosional forces, all transitions between a fully oxygenated and an anoxic sea are conceivable.

Sediment facies in the shallow water environment will mirror these events through distinct cycles and fossil assemblages. At O_2 times, the sediment distribution in the deep sea will resemble that of today, whereas at H_2S times, sapropels will occur, and calcareous oozes and marls will show pulses of dissolution. This process can go so far that a hiatus will develop and the dissolved chemical species will accumulate in the deep waters. Submarine volcanic activity and rifting will enhance these actions. It is of note that brines can accumulate in the deepest part of a basin. Upwelling or overturning of water masses will cause the release of these substances, and phosphorites, carbonates, and Fe–Mn oxides and hydroxides are a few of the major phases that may come into existence.

A graph (Fig. 6) shows schematically the pathway of four critical elements through geological time. The modern concentration level for each of the four elements in the deep sea is shown at the far right. The main features are as follows:

(1) Fe^{2+} becomes oxidised in the euphotic zone. Lowering of the O_2–H_2S interface will release Fe^{2+} to the upper layer causing the formation of siderite, a series of Fe-oxides and hydroxides, and Fe-silicates[34]. In the anoxic sea, Fe^{2+} is extracted as sulphide or silicate (for example, chlorite). Towards the end of the Precambrian, the ocean water turned over, fully oxygenating the deep sea for the first time, but went back to the anoxic state during the Silurian. This sequence is repeated several times throughout geological time.

Fig. 5 Formation and evolution of stratified waters (see text).
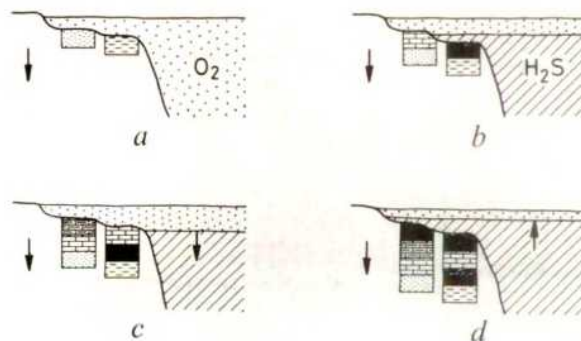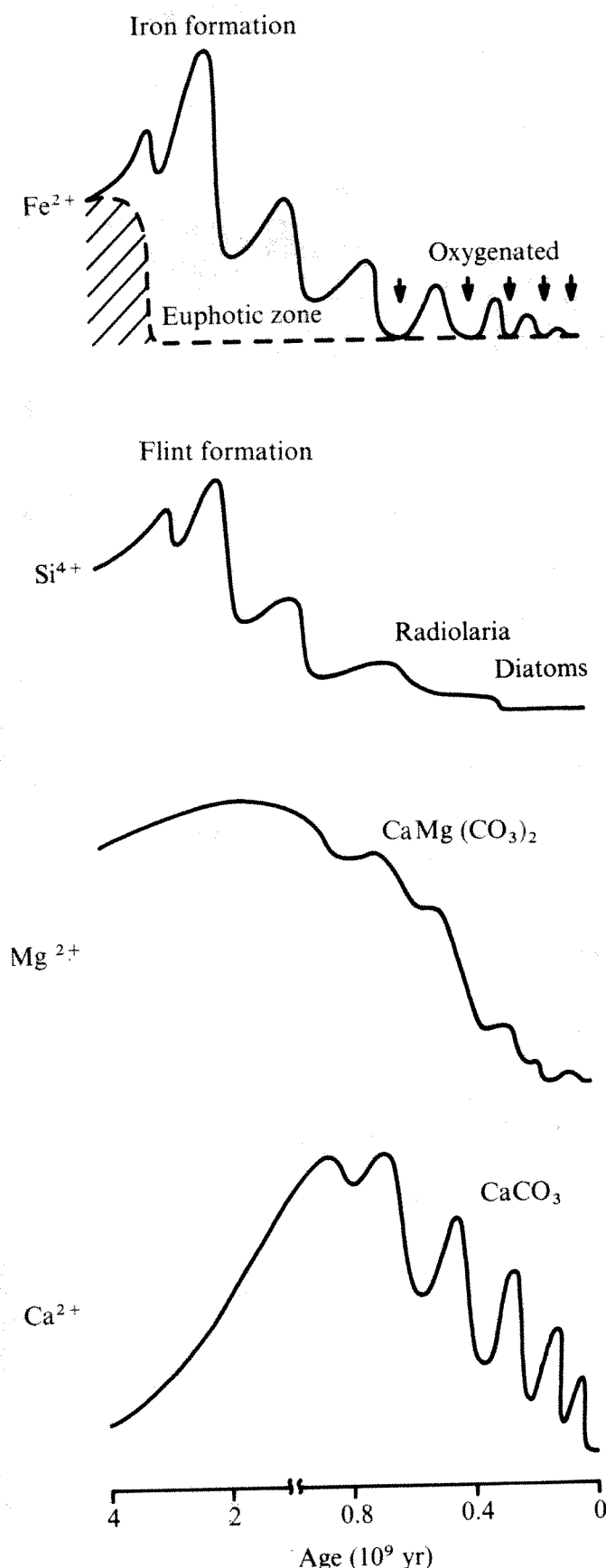
(2) In the Precambrian, the water is nearly saturated with silica below the interface. When released to the upper layer it precipitates on evaporation in shallow-water environments as amorphous silica. Biogenic silica extraction is the dominant release mechanism in more recent times.

(3) Mg^{2+} is extracted in the deep and shallow sea as silicate. Conditions for dolomite formation gradually develop and massive primary dolomite formation proceeds during the Palaeozoic and Mesozoic.

(4) Ca-carbonate deposition is dependent on the Mg–Ca ratio in the water phase[16]. Aragonite is the principal $CaCO_3$ species at the start of carbonate extraction, and high Mg-calcite, and eventually calcite, follow. Because of the instability of aragonite and high Mg-calcite, recrystallisation in the direction of hard dolomites and limestones is commonly observed in the older strata. In contrast, sediments composed of calcitic forams and/or coccoliths have a tendency to remain unindurated.

Conclusions

The highlight of the present study is the observation that permanently stratified modern marine and freshwater basins induce carbonate precipitation in the upper layer and carbonate dissolution in the anoxic deeper water. Vertical advection and diffusion will recycle dissolved chemical species back into the surface layer; however, a more substantial mineralisation of the euphotic zone is achieved by lowering the O_2–H_2S interface. This will promote massive carbonate precipitation. As a result sediments above a density boundary get rich in carbonates while those beneath have little or none. Because climate and/or geological settings change, thermo- and haloclines move up and down the water column or break up entirely, disrupting established ecosystems. The sediment record will reflect these incidents in the form of distinct cycles, facies, fossil assemblages, or hiatus.

Hutton's principle of uniformity of process, which holds that the present is the key to the past, gets an unexpected tenor. The well mixed oxygenated ocean of today seems not to be the model environment for the past 600 Myr. Instead, we are dealing with a sea which alternates between two opposing states: stratified and non-stratified. Depending on local or regional hydrographic settings, which may effect deep-water circulation and the hydrodynamics of water movement vertically and horizontally, one may encounter all sorts of transitions. At present, some areas in the open ocean are already anaerobic, and indicate the delicate balance between the oxic and anoxic states. All these findings lead us to conclude that a number of former interpretations regarding the origin of sedimentary cycles, the history of the oceans, and the workings of evolution have to be re-examined.

Fig. 6 Chemical evolution curves of Fe^{2+}, Si^{4+}, Mg^{2+}, and Ca^{2+} levels in the deep sea (schematic). The Fe^{2+} content in the upper layer drops to zero early in the Precambrian as soon as O_2-generation reaches a critical level.

1 Craig, H., *J. geophys. Res.*, **74**, 5491–5506 (1969).
2 Spencer, D. W., and Brewer, P. G., *J. geophys. Res.*, **76**, 5877–5892 (1971).
3 Brewer, P. G., and Spencer, D. W., *Am. Ass. Petrol. Geol., Mem.*, **20**, 137–143 (1974).
4 Riley, G. A., Stommel, W., and Bumpus, D. F., *Bull. Bingham oceanogr. Coll.*, **12**, 1–169 (1949).
5 Radach, G., and Maier-Reimer, E., *Proc. Sixth Coll. Ocean Hydrodynamics, Liege 1974* (in the press).
6 Kraus, E. B., and Turner, J. S., *Tellus*, **19**, 98–106 (1967).
7 Hasselmann, K., *et al.*, *Ergänz Hft Dt.-hydrogr. Ztschr., Reihe A* (8°), **12** (1973).
8 Kitaigorodskii, S. A., and Miropol'skii, Yu. Z., *Izv. Akad. Sci. USSR, atmos. Ocean Phys.*, **6**, 97–102 (1970).
9 Denman, K. L., *J. phys. Oceanogr.*, **3**, 173–184 (1973).
10 Degens, E. T., Von Herzen, R. P., Wong, H. K., Deuser, W. G., and Jannasch, H. W., *Geol. Rdsch.*, **62**, 245–277 (1973).
11 Degens, E. T., Von Herzen, R. P., and Wong, H. K., *Naturwissenschaften*, **58**, 229–240 (1971).
12 Müller, G., and Förstner, U., *Miner. Deposita*, **8**, 278–290 (1973).
13 Stoffers, P., *Sedimentpetrographische, geochemische und paläoklimatische Untersuchungen an ostafrikanischen Seen, Habilitationsschrift*, 117 (Universität Heidelberg, 1975).
14 Degens, E. T., and Hecky, R. E., *Colloqu. int. C.N.R.S.*, **219**, 13–24 (1974).
15 Hecky, R. E., and Degens, E. T., *Tech. Rep. Woods Hole Oceanogr. Inst., WHOI-73-28*, 93 (1973).
16 Müller, G., Irion, G., and Förstner, U., *Naturwissenschaften*, **59**, 158–164 (1972).
17 Verbeke, J., *Exploration Hydrobiologique des Lacs Kivu, Edouard et Albert*, **3**, Fasc. 1, 177 (1957).
18 Hustedt, F., *Exploration Parc Albert, Miss. H. Damas* (1935–1936), Fasc. 8, 199 (1949).
19 Coulter, G. W., *Limnol. Oceanogr.*, **8**, 463–477 (1963).
20 Beauchamp, R. S. A., *Int. Rev. Hydrobiol.*, **39**, 316–353 (1939).
21 Livingstone, D. A., *Limnol. Oceanogr.*, **10**, 607–610 (1965).
22 Talling, J. F., *Limnol. Oceanogr.*, **8**, 68–78 (1963).

[23] Brooks, J. L., Q. Rev. Biol., 25, 30–60, 131–176 (1950).
[24] Ladiges, W., Int. Rev. Hydrobiol., 53, 339 (1968).
[25] Moore, J. E. S., The Tanganyika Problem (Hurst and Blackett, London, 1903).
[26] Günther, R. S., Q. Jl Microsc. Sci., 36, 271 (1894).
[27] Degens, E. T., and Ross, D. A., Am. Ass. Petrol. Geol., Memoir, 20, (1974).
[28] Deuser, W. G., Am. Ass. Petrol. Geol., Mem., 20, 133–136 (1974).
[29] Redfield, A. C., Ketchum, B. H., and Richards, F. A., The Sea, 2 (edit. by Hill, M. N.), 26–77 (Interscience, London, 1963).

[30] Deuser, W. G., Deep-Sea Res., 18, 995–1004 (1971).
[31] Degens, E. T., Michaelis, W., Mopper, K., and Kempe, St., Neues Jb. Geol. Paläout., Mh. (in the press).
[32] Degens, E. T., and Ross, D. A., (eds), Hot Brines and Recent Heavy Metal Deposits in the Red Sea (Springer, New York, 1969).
[33] Plummer, L. N., and Wigley, T. M. L., Geochim. cosmochim. Acta, 40, 191–202 (1976).
[34] Floran, R. J., and Papike, J. J., Geol. Soc. Am. Bull., 86, 1169–1190 (1975).

Similarities and differences in the structure of X and Y chromosome rRNA genes of *Drosophila*

Kenneth D. Tartof
Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Igor B. Dawid
Department of Embryology, Carnegie Institution of Washington, Baltimore, Maryland 21210

In Drosophila melanogaster, *the rRNA genes (rDNA) are clustered at single sites on two non-homologous chromosomes, the X and Y. Examination of the structure of X and Y rDNAs with restriction endonucleases reveals that the X rDNA contains repeating units not present in the Y. Such obervations, as well as genetic evidence, illustrate difficulties with the hypothesis that recombination is the predominant mechanism preserving similarities between these two clusters maintained on different chromosomes. This raises the possibility that selection pressure has a significant role in maintaining the parallel evolution of these two separate but homologous redundant gene clusters.*

In the fruit fly, *Drosophila melanogaster*, there is one nucleolus organiser (*NO*) on the X chromosome and one on the Y. At each *NO*, there are 200–250 tandemly arranged ribosomal RNA genes (rDNA)[1,2]. Although these genes reside on non-homologous chromosomes, the 18S and 28S RNA molecules transcribed from either the X or Y rDNA have very similar, and probably identical, primary sequences as revealed by fingerprinting experiments.[3] Furthermore, in the wild-type *D. melanogaster* male, which is genotypically XY, there is no meiotic recombination[4] though mitotic recombination occurs in both both sexes[5]. Thus not only are the rRNA genes located on non-homologous chromosomes but they also seem to be genetically isolated from one another by virtue of the absence of meiotic recombination.

Attempts to elucidate the detailed molecular structure of *Drosophila* rDNA by electron microscopic analysis of actively transcribing clusters of rRNA genes have produced conflicting results. Hamkalo *et al.*[6] found a very heterogeneous rDNA unit repeat length when embryonic material was used. Using similar techniques, Laird and Chooi (personal communication) have found that two length classes of rDNA repeats exist, both of which occur on the X *NO*, but only the shorter one was observed on the Y *NO*. Observations on rDNA fragments produced by the restriction endonuclease *Eco*RI and cloned in *Escherichia coli* revealed two distinct size classes of repeating units, one 17 kb (kilobases or kilobase pairs)[7] and the other, 11–12 kb (D. M. Glover and D. S. Hogness, personal communication). We report here the distribution of these size classes on the X and Y *NO*. Our results define the restriction endonuclease products of the X and Y rDNA, and demonstrate that both *NOs* are similar with respect to one major class of repeating

elements but differ decisively in their content of another class of repeats.

Restriction endonuclease analysis of X and Y chromosome rDNA

On treatment of unfractionated DNA of appropriate genetic origin with *Eco*RI the fragments were separated by Agarose gel electrophoresis and transferred to nitrocellulose filters according to the method of Southern[8]. The filter-bound fragments were hybridised to radioactively-labelled rRNA and finally made visible by autoradiography. Figure 1*a* shows the results of a series of experiments in which females (X/X) from two different wild-type strains (Ore-R and Lausanne) and males carrying only the Y chromosome *NO* are compared. It is obvious that the X chromosome contains both 17- and 11-kb segments while the Y contains the 11-kb but not the 17-kb fragment. In addition to these bands, two components of about 7.4 kb and 5.4 kb occur in both X and Y *NOs*. A mixture of X and Y DNA digested with *Eco*RI also displays these four bands, indicating that there is no apparent difference in the size of *Eco*RI fragments derived from X or Y chromosomes. When purified rDNA is digested with *Eco*RI there are, in addition to the predominant 17-, 11-, 7- and 5-kb size classes, minor fragment classes of intermediate length, some of which can be seen as bands in Fig. 1*b*. Somewhat similar length heterogeneities have been observed among *Xenopus* rDNA repeats[12].

From mapping studies it is known that the 11-kb and 17-kb fragments represent complete repeating units of the rDNA, which are distinguishable by the insertion of an additional spacer sequence within the 28S gene region (Glover and Hogness, personal communication; unpublished results of P. K. Wellauer and I.B.D.). The nature of the shorter 5- 7-kb fragments is not yet analysed, but it seems likely that they arise from the presence of additional *Eco*RI sites in a fraction of the repeating units. The other minor bands are due to additional *Eco*RI sites and to the length heterogeneity in the repeating unit.

The difference between the X and Y *NOs* can be demonstrated by other means. The restriction enzyme from *Serratia marscencens, Sma*I, yields a different fragment pattern from the two types of rDNA (Fig. 2). Y chromosome rDNA yields one fragment about 8.5 kb long and a second fragment 2.8 kb long, both of which hybridise to 28S and 18S RNA. *Sma*I digestion of X rDNA, however, produces a further fragment of 2.2 kb which hybridises with 28S but not with 18S RNA (Fig. 2). Another fragment of 2.3 kb can be seen in the ethidium bromide stained pattern obtained with rDNA purified from a mixture of males and
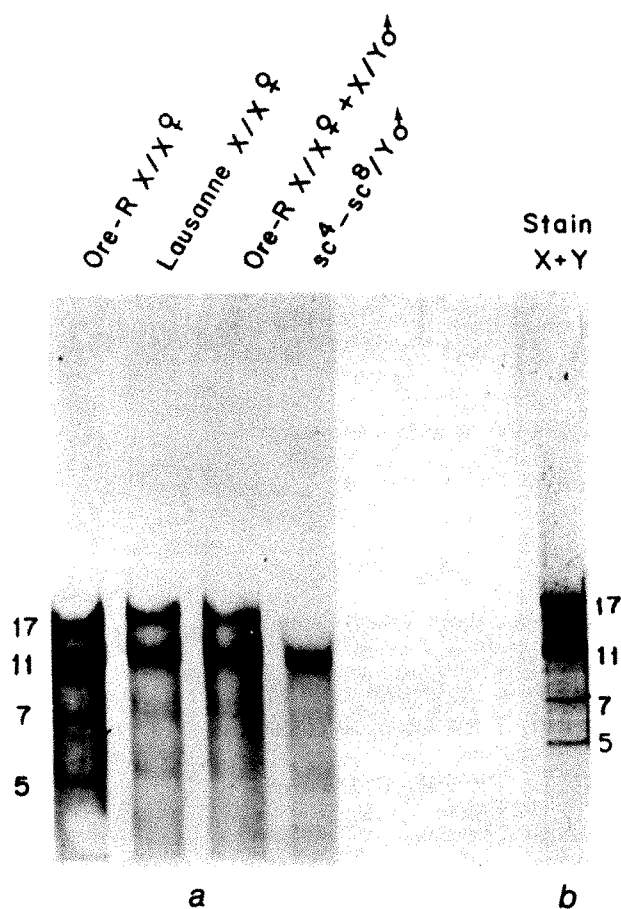
Fig. 1 *Eco*RI restriction pattern of *Drosophila* rDNA. The enzyme was prepared according to Greene *et al.*[9]. DNA of the indicated genotype was extracted, digested with *Eco*RI endonuclease and subjected to electrophoresis in 1% Agarose for 4 h at 100 V. The DNA fragments were transferred to a nitrocellulose filter[7] which was then hybridised to ^{125}I-labelled rRNA, 5×10^7 c.p.m. μg^{-1}, for 24 h in $2 \times$SSC at 60 °C. The filter was treated with RNase (30 μg ml^{-1}) at 37 °C for 60 min, dried and finally exposed to X-ray film for 5–10 d. The size of each fragment class is indicated in kb and was determined using *Eco*RI-treated λ DNA as the standard[10]. *a*, Ore-R wild-type female, Lausanne wild-type female, Ore-R female + male, and *sc⁴–sc⁸/Y* (Ore-R) male (*sc⁴–sc⁸* designates an X chromosome that is deficient for its *NO*, see ref. 11); *b*, purified rDNA from a mixture of males and females that has been digested with *Eco*RI with the resulting fragments separated by electrophoresis and then stained with ethidium bromide.

females, but it does not hybridise with rRNA and does not appear in the autoradiogram as do the other fragments. Thus the analysis with *Sma*I again confirms the presence of a structural difference between X and Y chromosome rDNA.

Fragment patterns of rDNA from *bobbed* mutants

To gain some insight into the relative interspersion of 17- and 11-kb units among each other we also analysed the fragment pattern obtained with *Eco*RI, using rDNA from four independently isolated X chromosome *bobbed* mutations. These mutations are deleted for 50–60% of their rDNA. If the NO consists principally of clusters of 17- and 11-kb units segregated from each other, then mutants deleted for 50% of their rDNA would be expected to be under represented for one class of the repeated units. Figure 2*b* shows that the rDNA in these mutants contains both the 17- and 11-kb size classes and that they are in roughly the same proportion to each other as in the wild type. Similar results have been obtained with six other independently

derived *bobbed* mutants. Because all these *bobbed* mutants arose by the deletion of half the rDNA repeats from an original *NO* and because each deletion arose independently, the results argue for interspersion of 11-kb and 17-kb repeat lengths. It is interesting that conspicuous length heterogeneity also occurs in the rDNA of most *bb* mutants (except *bb⁶*). The *bb⁸¹* mutant is actually a doublet in the 11-kb region. One band is 11 kb, the other 10 kb. *bb³* is extremely heterogeneous in the range 11 to 9 kb, revealing no discrete size class at the resolution of our gel, while *bb⁴* contains bands at 11 and 9 kb and many smaller bands. Such extensive heterogeneity has been observed consistently among *bobbed* mutants. In wild-type rDNA length heterogeneity does exist, but most repeating units occur as the 11- and 17-kb size classes (Fig. 1*b*). One possible explanation for this heterogeneity is that *bobbed* mutants result in the preferential retention of some minor size classes during deletion of predominantly 11- and 17-kb repeats. The heterogeneity of lengths of *Eco*RI fragments in *bb⁴*, however, are so extensive that they cannot be explained in this way and we must consider additional mechanisms for generating such variations. Perhaps this length heterogeneity is a consequence of the event that initially generated the *bb⁴* partial deletion.

Fig. 2 Fragments obtained from *Drosophila* rDNA with *Sma*I. The restriction enzyme was purified and used according to C. Mulder (personal communication). Fragments were separated and hybridised as described for Fig. 1. The stained pattern (ethidium bromide, 1 μg ml^{-1}) was obtained from purified rDNA (X/X females and X/Y*bb* males; because the X contains about 200 rRNA genes and the Y*bb* about 120, 83% of the rDNA of this stock is derived from the X). The purification of the rDNA using gradient centrifugation in the presence of mercury and actinomycin D will be published elsewhere. The X and Y chromosome rDNA was partially purified before digestion, electrophoresis and hybridisation. Each filter lane was cut in half lengthwise and the left half was hybridised with 28S RNA, the right half with 18S RNA. The length of different fragments is given in kb.
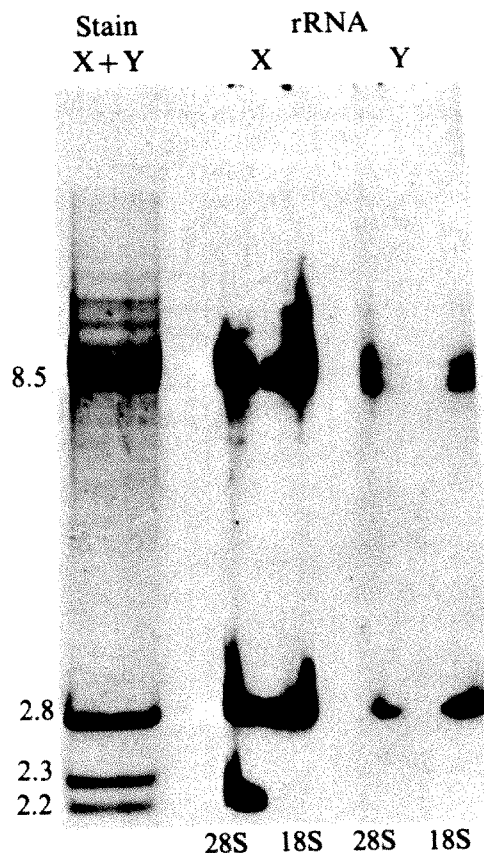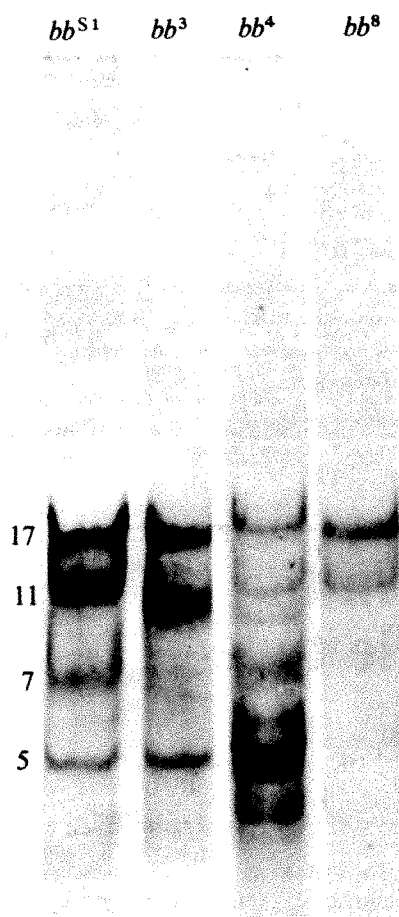
bb^{S1} bb^3 bb^4 bb^8

Fig. 3 *Eco*RI restriction pattern of rDNA from *bobbed* mutants. DNAs from females homozygous for bb^{S1}, bb^3, bb^4 and bb^8 mutations were extracted, treated with *Eco*RI and analysed by electrophoresis and hybridisation to ^{125}I-rRNA as described in Fig. 1.

Genetic mechanisms in the evolution of rDNA

The results of the experiments described here, and summarised in Fig. 4, have interesting implications for the mechanism of redundant gene evolution. If the X and Y *NO*s were to recombine with each other (and these would necessarily have to be double exchanges to maintain intact X and Y chromosomes; Fig. 4), and different repeat lengths are interspersed, then both *NO*s would be expected to display similar *Eco*RI fragment patterns. The fact that the 17-kb fragment is present on the X but absent from the Y chromosome supports earlier genetic evidence against recombination between X and Y rDNAs. It then becomes necessary to explain the fact that both the X and Y *NO*s contain apparently identical 18S and 28S RNA sequences[3], and share three of the four major discrete size classes of *Eco*RI fragments and at least one discrete fragment (2.8 kb) of the *Sma*I digest (Fig. 4). Unless genetic exchange between the X and Y *NO*s occurs on some occasions, we conclude that the similarities in the rDNA of these two *NO*s are maintained by selection pressure.

Drosophila is not unique in maintaining homologous genes on more than one chromosome. In humans, five chromosomes each contain a single NO[14,15]. In various organisms the same satellite DNA sequence is frequently found to cluster at the centric heterochromatic regions on several non-homologous chromosomes[16]. In *Xenopus*, the 5S RNA genes are located at the telomeres of at least 15 chromosomes[17]. These are but a few examples which illustrate the variety of homologous genetic sites that have been found to

be dispersed throughout the genome of a given species. In the case of the *Xenopus* 5S genes, it has been suggested that the relative homogeneity among the 15 gene clusters has been maintained by recombination[17,18].

In the examples cited above it has not been tested whether the homologous genes or satellites found on different chromosomes are truly identical or differ in some structural aspects. Furthermore, although there is some cytological evidence for recombination between some of these homologous regions on non-homologous chromosomes, as in the case of *Xenopus* 5S DNA[18], there is no direct genetic or physical evidence for such exchanges.

The situation with *Drosophila* rDNA poses a dilemma, in that two opposing interpretations of the available evidence may be proposed, both of which necessitate certain assumptions. The selection hypothesis is based on the prior genetic evidence against X–Y crossover events and places weight on the absence of 17-kb rDNA fragments on the Y *NO*. With these points in mind, we conclude that the X and Y *NO*s are indeed genetically isolated from each other. The apparent identity of the primary structure of the 18S and 28S rRNAs[3], and the similarity of the restriction fragment patterns which we observe would then be due to selection pressures. Although it is difficult to understand how selection would be sensitive to details of the spacer structure and to every nucleotide in the rRNA sequence, it must be remembered that the functional significance of spacer DNA and all of the structural gene information is still obscure. On the other hand, the recombination hypothesis asserts that genetic exchanges in the form of double crossing-over events do take place between the X and Y *NO*s. Such exchanges would not be detected by genetic experiments, and they would account for those similarities between the X and Y rDNAs which exist. This interpretation, however, necessitates the assumption that rDNA repeats of the 17-kb class

Fig. 4 A summary of the similarities and differences in the structure of X and Y rRNA genes. Since one-third of the X DNA resides in the heterochromatic block[13] (indicated as the darkened portion of the X) and since there are about 250 rRNA genes per X NO^2, this means that the X rDNA accounts for about 30% of DNA in the heterochromatin of this chromosome. Similarly, about 30% of the DNA in the short arm of the Y is rDNA. *Eco*RI and *Sma*I fragment sizes are given in kb. Infrequent *Eco*RI size classes such as those seen in Fig. 1b have not been included. Parentheses around certain size classes indicate they are minor components.

*Eco*RI	*Sma*I
17	8.5
11	2.8
(7.4)	2.3
(5.4)	2.2

rDNA

*Eco*RI	*Sma*I
11	8.5
(7.4)	2.8
(5.4)	2.3
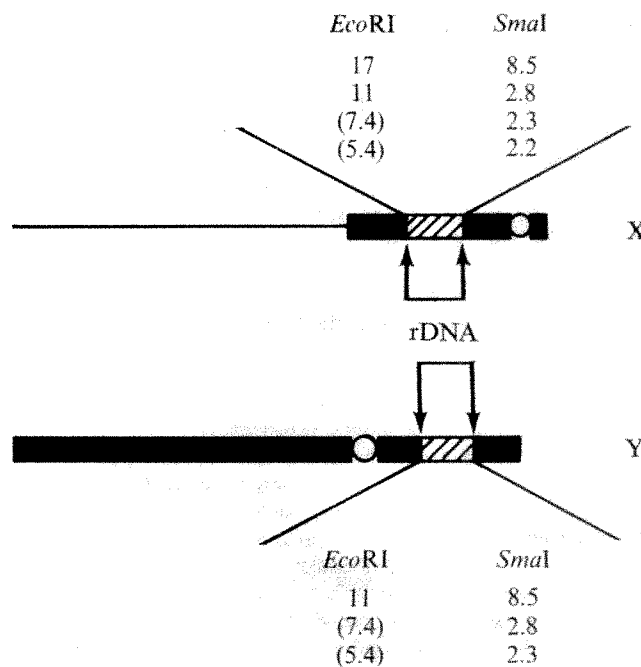
either cannot participate in genetic exchange or are rapidly eliminated from the Y chromosome *NO*. This is a purely *ad hoc* assumption without any supporting experimental basis.

Our obsevations do not allow an unequivocal conclusion on the genetic and evolutionary mechanisms which determine the structure of *Drosophilia* rDNA, and which may operate in clusters of repetitive genes in general. Nevertheless, this work points up difficulties with the hypothesis that genetic exchanges are the main mechanism which assures the similarity of homologous regions on non-homologous chromosomes, and raises the possibility that selection pressures play a significant role in the parallel evolution of families of repetitive genes.

1 Ritossa, F. M., Atwood, K. C., Lindsley, D. L., and Spiegelman, S., *Natn. Cancer Inst. Monog.*, **23**, 449–472 (1966).
2 Tartof, K. D., *Science*, **171**, 294–297 (1971).
3 Maden, B. E. H., and Tartof, K., *J. molec. Biol.*, **90**, 59–64 (1974).
4 Morgan, T. H., *Biol. Bull.*, **26**, 195–204 (1914).
5 Stern, C., *Genetics*, **21**, 625–730 (1936).
6 Hamkalo, B. A., Miller, O. L., and Bakken, A. H., *Cold Spring Harb. Symp. quant. Biol.*, **38**, 915–919 (1973).
7 Glover, D. M., White, R. L., Finnigan, D. J., and Hogness, D. S., *Cell*, **5**, 149–157 (1975).
8 Southern, E. M., *J. molec. Biol.*, **98**, 503–517 (1975).
9 Greene, P. J., Betlach, M. C., Boyer, H. W., and Goodman, H. M., *Meth. molec. Biol.*, **7**, 87–111 (1974).
10 Allet, B., Jeppesen, P. G. N., Katagui, K. J., and Delius, H., *Nature*, **241**, 120–123 (1973).
11 Lindsley, D. L., and Grell, E. H., *Carnegie Inst. Wash. Publ.*, **627** (1968).
12 Wellauer, P. K., *et al.*, *Proc. natn Acad. Sci. U.S.A.*, **71**, 2823–27 (1974).
13 Rudkin, G. T., *Genetics Today, Proc. eleventh Int. Cong. Genet.*, 359–374 (1963).
14 Henderson, A. S., Warburton, D., and Atwood, K. C., *Proc. natn Acad. Sci. U.S.A.*, **69**, 3394–3398 (1972).
15 Evans, H. J., Buckland, R. A., and Pardue, M. L., *Chromosoma* (in the press).
16 Tartof, K. D., *A. Rev. Genet.*, **9**, 355–385 (1975).
17 Brown, D. D., and Sugimoto, K., *J. molec. Biol.*, **78**, 397–415 (1973).
18 Pardue, M. L., Brown, D. D., and Birnstiel, M. L., *Chromosoma*, **42**, 191–203 (1973).

Significance of impulse activity in the transformation of skeletal muscle type

S. Salmons and F. A. Sréter

Department of Anatomy, Medical School, University of Birmingham, Birmingham B15 2TJ, UK, and Department of Muscle Research, Boston Biomedical Research Institute, 20 Staniford Street, Boston, Massachusetts 02114

The changes which follow cross reinnervation of mammalian fast and slow twitch muscles may reflect a capacity of skeletal muscle to respond adaptively to different functional requirements. This interpretation is supported by experiments in which long-term electrical stimulation was used both to reproduce and to oppose the effects of cross reinnervation.

IN adult mammals the motoneurones innervating slow twitch skeletal muscles generate a sustained low-frequency pattern of activity; in contrast, motoneurones supplying fast twitch muscles subject them to intermittent bursts of more intense activity[1]. The characteristic physiological and biochemical properties of the muscles emerge late in development, usually after birth, and in slow muscles this process seems to depend on the establishment of the adult pattern of motoneurone activity. If the motoneurones are rendered quiescent by isolation of the spinal cord a few days after birth, muscles which would normally have acquired slow contractile characteristics develop as fast muscles[2]. Even in the adult animal a fast muscle subjected to low frequency activity by an implanted stimulator[3] undergoes profound changes in all the properties which normally distinguish it from a slow muscle[4–8]. It has been pointed out[4,5] that this capacity for change would enable a muscle to accommodate the changing functional requirements which may be encountered during an animal's lifetime. According to this view, the late phase of differentiation of slow muscles in the neonate and the changes resulting from chronic stimulation of fast muscles in the adult both represent an adaptive response to continuous, low-frequency activation. The question then arises as to whether such a concept can provide insight into previous demonstrations of plasticity of skeletal muscle properties. The most important of these are the cross reinnervation experiments pioneered by Buller *et al.*[9].

In these experiments the motor nerves of a fast and a slow muscle were cut and cross anastomosed; as a result, the fast muscle was reinnervated by the nerve which previously supplied the slow muscle, and vice versa. The contractile characteristics then changed, the fast muscle becoming slower and the slow muscle faster. The authors considered, but were not convinced by, the possibility that these changes were due to changes in the patterns of impulse activity reaching the muscles. Instead they proposed that the contrasting properties of slow and fast muscles were

Fig. 1 Isometric twitch contractions of left and right EDL muscles (*a*) and right soleus muscle (*b*) recorded on identical time scales. The left EDL muscle (slower time course) has been stimulated at 10 Hz continuously for 20 weeks. Each time scale represents 100 ms with 1-ms divisions. Peak tensions are given in Table 1.
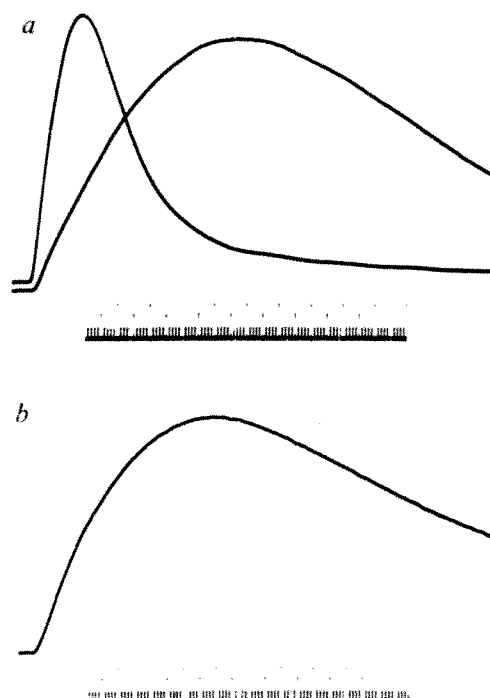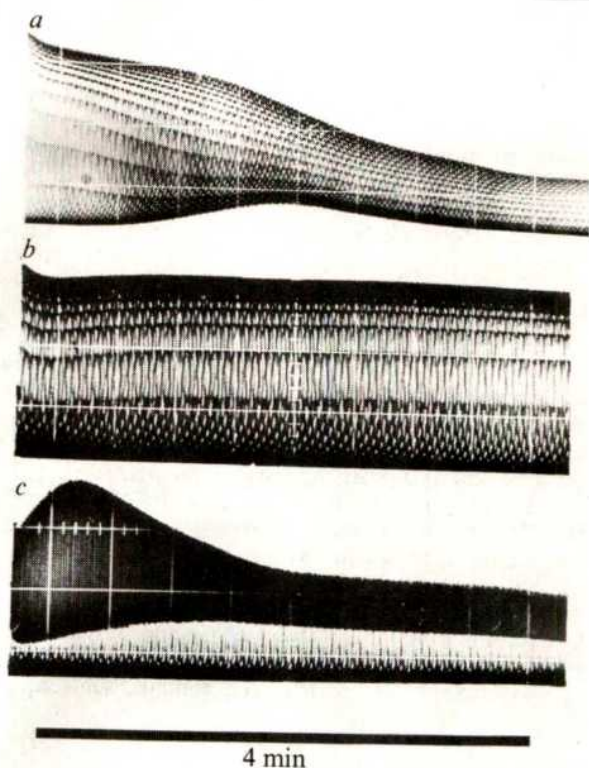
Fig. 2 Response of control soleus (*a*), stimulated EDL (*b*) and control EDL (*c*) muscles to a prolonged train of unfused tetani (25 Hz for 500 ms repeated every 1.25 s). The stimulated muscle is remarkably resistant to fatigue in these taxing conditions.

attributable to chemical trophic influences exerted by their respective motor nerves; cross union of the nerves caused the trophic substances to be carried to muscles of the opposite type, which were then transformed in accordance with the new innervation. Trophic concepts were not foreign to developmental biology, and the hypothesis was widely supported.

If cross union of the nerves interchanges the characteristic patterns of impulse activity reaching the slow and fast muscles, changes of the type observed would be in accordance with the adaptive-response hypothesis. The fast muscles, subjected to a continuous low frequency discharge, would tend to acquire slow muscle properties; the slow muscles, relieved of such a pattern of activation, would tend to undergo reciprocal changes. Although it remains to be confirmed that the firing pattern of a given moto-

neurone pool is fundamentally the same after cross reinnervation as before, Yellin[10] has reviewed indirect evidence which strongly supports this expectation.

Assuming, therefore, that there is a transposition of activity, can this alone account for the effects of cross reinnervation? This question has received balanced discussion (for example, refs 11 and 12) but has remained unanswered because of a lack of suitable data. We wish to report new evidence which makes it possible to compare, both qualitatively and quantitatively, the effects of cross reinnervation with those of a change in the pattern of impulse activity.

Long term stimulation

In their original work Salmons and Vrbová demonstrated marked slowing in response to long term stimulation in fast muscles of the rabbit and cat[5]. The results were confirmed and extended by two other groups, using different techniques of stimulation[13,14]. These experiments showed that the effects of stimulation were significant and were not confined to a particular muscle or species, but the duration of stimulation was not sufficient to afford a basis for comparison with cross reinnervation experiments. As a result of improvements in the design of the stimulator and in the surgical technique for its implantation, we have been able to stimulate fast muscles of the rabbit continuously for much longer periods than before. We present here the results of two experiments, each of 20 weeks' duration.

The animals were operated on in strictly aseptic conditions. In each case the stimulator capsule was implanted intraperitoneally, and the two flexible leads were passed subcutaneously to the left hind limb and fixed near the common peroneal nerve. The right limb served as an unoperated control. The animals were returned to the colony for the ensuing period, during which they remained in excellent condition and showed a normal gain in body weight. After 20 weeks a terminal procedure was carried out in which isometric twitch, tetanus and fatigue characteristics were determined for the tibialis anterior (TA), extensor digitorum longus (EDL) and soleus muscles of both hind limbs. Each muscle was set at the resting length at which it produced the maximum twitch contraction. Recordings from corresponding muscles of the two sides were made simultaneously in controlled thermal conditions. The muscles were then quickly excised, weighed, frozen in liquid nitrogen, sealed hermetically in polyethylene sachets and stored in solid CO_2 for biochemical analysis.

Figure 1 compares the twitch contraction of a stimulated EDL muscle with that of control EDL and soleus muscles of the contralateral limb. It is evident that the stimulated EDL muscle contracts even more slowly than a classical

	Control TA	Stimulated TA	Control soleus
Table 1 Effects of long term stimulation on a fast muscle			
Maximum twitch tension (g)	370	455	251
Time to peak twitch contraction (ms)	17.0	64.9	57.2
Time from peak to half-relaxation (ms)	17.9	77.1	91.1
Maximum tetanic tension P_0 (g)	3919	1829	1254
Twitch : tetanus ratio	0.094	0.249	0.200
Maximum rate of rise of tetanic tension ($\%P_0$ ms^{-1})	3.06	1.11	1.73
Ca^{2+}-activated myosin ATPase activity (μmol mg^{-1} min^{-1})	0.66	0.18	0.22
Ca^{2+}-activated myosin ATPase activity after preincubation at pH 9.2 (μmol mg^{-1} min^{-1})	0.68	0.06	0.06
Ca^{2+} uptake by FSR ; initial rate (μmol mg^{-1} min^{-1})	2.19	0.21	0.56
Total uptake in 15 min (μmol mg^{-1})	3.81	0.55	1.52

Physiological and biochemical characteristics of the muscles whose twitch contractions are illustrated in Fig. 1. Fragmented sarcoplasmic reticulum (FSR) and myosin were prepared as reported earlier[20,21]. To determine the alkali lability myosin was preincubated at pH 9.2 in a solution containing 0.2 mg of protein per ml, 25 mM KCl, 10 mM Tris, 10 mM CaCl$_2$ for 10 min at 25 °C. At the end of this preincubation period the pH was readjusted to 7.6 by adding Tris to a final concentration of 50 mM and HCl as required. The ATPase reaction was then started by adding ATP.
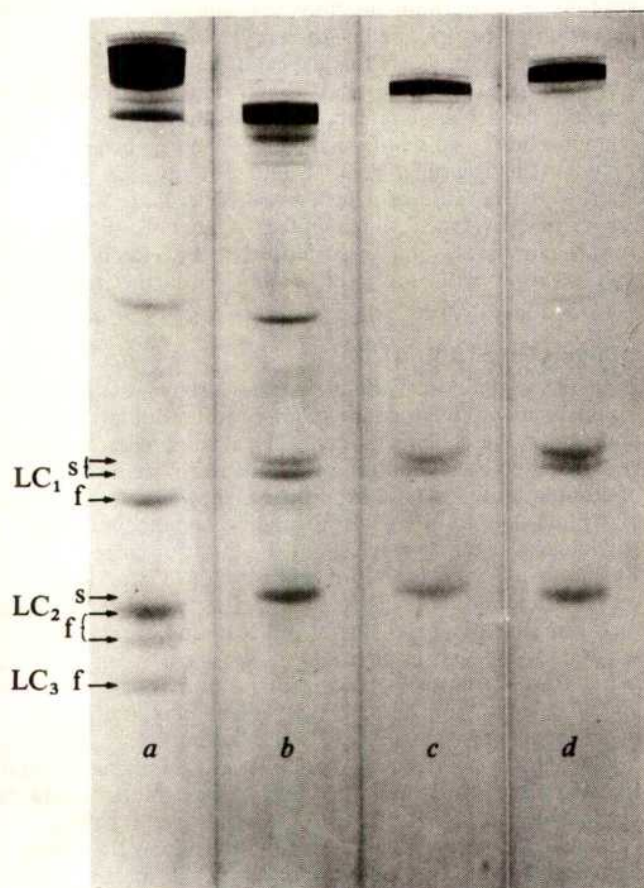
Fig. 3 SDS–polyacrylamide gel electrophoretograms of myosin prepared from control TA muscle (*a*) TA muscle stimulated for 20 weeks (*b*), EDL muscle stimulated for 20 weeks (*c*), control soleus muscle (*d*). Myosin light chains characteristic of slow muscle are demonstrable after 4 weeks of stimulation[15]. Continued stimulation results in the successive disappearance of the fast muscle light chains, in the order LC_{2f}, LC_{3f}, LC_{1f} (refs 6 and 8).

qualitatively and quantitatively for the changes which result from cross reinnervation.

Equality of effect, however, does not necessarily imply equivalence of mechanism. It could be argued that changes in impulse activity and cross union of motor nerves achieve similar end results by different routes. To examine this possibility, we designed a null experiment in which the effects of long term stimulation would be set in opposition to those of cross reinnervation (for a preliminary account see ref. 17).

Stimulation versus cross reinnervation

In a group of 20 rabbits the common peroneal nerve and the motor nerve to the soleus muscle were exposed using a lateral approach (A. R. Luff, personal communication). One of the branches of the deep peroneal nerve supplying motor fibres to the fast muscles TA and EDL was carefully dissected free, cut and gently apposed to the peripheral cut end of the soleus nerve, using a single suture of 9/0 or 10/0 nylon. The reciprocal cross union was not performed, to provide conditions which favoured cross reinnervation of soleus but militated against self reinnervation by its original nerve. Since the bulk of the common peroneal nerve was left intact, the functional impairment incurred by this operation was minimal. Eight weeks after cross union, a stimulator was implanted in each of ten animals, subjecting the

slow muscle in the same animal. Similar results were obtained for the other stimulated fast muscles. In response to a prolonged train of 25-Hz tetani, the stimulated muscles proved to be even more resistant to fatigue than the solei, and showed, in the early phase of the fatigue record, the slight post-tetanic depression characteristic of slow muscles in place of the marked potentiation normally seen in fast muscles (Fig. 2). The biochemical results were no less striking, and showed a continuation of the trend described previously[6]. Table 1 illustrates typical data from one muscle pair. In addition to changes in these parameters, electrophoretograms of myosin light chains on sodium dodecyl sulphate (SDS)–polyacrylamide gels revealed that the light chains characteristic of fast muscles had been replaced entirely by light chains typical of slow muscles (Fig. 3). There was also evidence that the heavy chains of myosin had been replaced, since myosin from the stimulated muscle lacked the N-τ-methylhistidine amino acid residue characteristic of fast muscle heavy chains. Both these observations agreed with our earlier findings for shorter periods of stimulation[6,8,15,16].

These results represent a degree of transformation which matches or exceeds that of muscles cross reinnervated for much longer periods. We have shown, for example, that stimulation even for 12 weeks is enough to bring about the replacement of light and heavy chains of myosin noted above, yet both processes are still incomplete 1 yr after cross reinnervation[8,15]. The effects of a change in the pattern of impulse activity are therefore sufficient to account both

Fig. 4 Isometric twitch contractions of soleus muscles. In each case the left muscle of the pair has been reinnervated by the motor nerve from a fast muscle as a result of nerve cross union 16 weeks previously. In the upper picture the left muscle shows the increase in contractile speed which is the well documented consequence of cross reinnervation. The left muscle in the lower picture had been subjected to tonic stimulation through the new innervation for the previous 8 weeks; it exhibits a time course of contraction slightly more prolonged than that of the control muscle. Contractions of the left soleus muscles were elicited by stimulation of the left common peroneal nerve. In this and all other experiments of the series evidence for self reinnervation was sought by stimulating the left tibial nerve. The tension record then showed a small excursion of irregular shape and rapid time course (time to peak 20.07 ± 2.30 ms) which we attribute to residual mechanical cross-talk from the subjacent flexor hallucis longus muscle. In no instance was there any evidence of the slow twitch contraction which would have been indicative of self reinnervation.

Table 2 Characteristics of cross-reinnervated soleus muscles subjected to stimulation

	(a) Cross innervation alone		(b) Controls		(c) Cross innervation & stimulation
Time to peak twitch contraction (ms)	25.7 ± 0.7 $(n=6)$	$(P < 0.001)$	55.5 ± 2.9 $(n=16)$	$(P > 0.2)$	63.5 ± 8.3 $(n=6)$
Time to half-relaxation (ms)	26.7 ± 1.6 $(n=6)$	$(P < 0.001)$	74.2 ± 5.0 $(n=16)$	$(P > 0.9)$	72.9 ± 13.7 $(n=6)$
Twitch–tetanus ratio	0.15 ± 0.02 $(n=5)$	$(P < 0.01^*)$	0.21 ± 0.01 $(n=15)$	$(P > 0.2^*)$	0.16 ± 0.03 $(n=4)$
Max. rate of rise ($\%P_0 ms^{-1}$)	2.86 ± 0.17 $(n=5)$	$(P < 0.001)$	1.74 ± 0.14 $(n=13)$	$(P > 0.2)$	1.36 ± 0.21 $(n=4)$
Ca^{2+}–activated ATPase ($\mu mol\ mg^{-1}\ min^{-1}$)	0.48 ± 0.06 $(n=6)$	$(P < 0.001)$	0.26 ± 0.02 $(n=27)$	$(P > 0.6)$	0.28 ± 0.04 $(n=5)$
Alkali lability (% loss ATPase activity)	12.4 ± 6.3 $(n=6)$	$(P < 0.001)$	60.6 ± 2.8 $(n=15)$	$(P > 0.5)$	57.7 ± 2.6 $(n=5)$
Ca^{2+} uptake by FSR: initial rate ($\mu mol\ mg^{-1}\ min^{-1}$)	1.42 ± 0.24 $(n=6)$	$(P < 0.01)$	0.64 ± 0.11 $(n=17)$	$(P > 0.2)$	0.36 ± 0.09 $(n=5)$
total uptake ($\mu mol\ mg^{-1}$ per 15 min)	2.19 ± 0.10 $(n=6)$	$(P < 0.001)$	1.33 ± 0.09 $(n=17)$	$(P > 0.8)$	1.15 ± 0.31 $(n=5)$
ATPase ($\mu mol\ mg^{-1}\ min^{-1}$)	1.68 ± 0.29 $(n=6)$	$(P < 0.05)$	1.14 ± 0.12 $(n=17)$	$(P > 0.2)$	0.81 ± 0.17 $(n=5)$

*Paired *t* test.

Physiological and biochemical characteristics (mean±s.e.m.) are given for: *a*, soleus muscles cross reinnervated by a fast muscle nerve; *c*, muscles treated similarly but subjected to chronic stimulation during the last 8 weeks of the 16-week experimental period, and *b*, control soleus muscles. The data in columns (*a*) and (*c*) have been compared with control data in column (*b*) and *P* values for the differences assigned in each case. Except where noted, significance was assessed using Student's *t* test on the unpaired data. Additional control muscles have been included to establish reliable baselines and to provide more stringent conditions for the test of non-significance on the right hand side. In spite of the low mean value for the twitch–tetanus ratio in column (*c*), two of the animals in this group showed a higher twitch–tetanus ratio for the experimental muscle than for the contralateral control muscle. Of the muscles subjected to cross reinnervation alone (column (*a*)), none had a twitch–tetanus ratio greater than 60% of that of the contralateral control muscle.

cross reinnervated soleus muscle to a continuous train of impulses at 10 Hz delivered through the new innervation. These muscles would therefore have a pattern of activation substantially similar to that which preceded cross innervation. In the remaining animals dummy electrodes were implanted which were not connected to a stimulator. The rabbits were operated in matched pairs so that the two groups comprised animals with comparable histories. After a further 8 weeks the muscles were examined in a terminal procedure similar to that described above.

The adaptive response hypothesis required that the usual consequences of cross innervation be totally absent in the muscles which received stimulation. This expectation was indeed borne out. In Fig. 4 twitch contractions of soleus muscles from one pair of animals are compared. In the animal with dummy electrodes the cross innervated soleus muscle shows the usual substantial increase in the speed of contraction and relaxation. In the animal with a stimulator the cross innervated and stimulated soleus muscle contracts and relaxes even more slowly than its counterpart on the control side. Table 2 gives values for a broad range of physiological and biochemical variables measured in the two groups of animals. In each parameter the cross reinnervated soleus muscles differed significantly from the control muscles, whereas the muscles which were both cross reinnervated and stimulated did not differ significantly from the controls. Gel electrophoretograms of myosin light chains confirmed the presence, in cross reinnervated soleus muscles, of light chains characteristic of fast muscle, as reported previously[18]; on the other hand, cross reinnervated soleus muscles. Since it is highly unlikely that different chain patterns indistinguishable from those of control soleus muscles. Since it is highly unlikely that different mechanisms could produce such a detailed complementarity of effect we conclude that the effects of impulse activity and those of cross reinnervation are mediated in the same way.

A converse experiment would be to cross reinnervate a fast muscle with a nerve which previously supplied a slow muscle, and then relieved the muscle of the activity imposed by the new innervation. In fact this experiment formed part of the original study by Buller *et al.*[9]. No effects attributable to cross reinnervation could be discerned when steps had been taken to abolish normal motor activity by section of the spinal cord and lumbosacral dorsal roots. It was noted at the time that this observation could be interpreted within the framework of the trophic hypothesis if the synthesis, transport or release of trophic substances were dependent on the passage of nervous impulses. The finding is, however, equally well understood in terms of a more direct role of impulse activity, and as such is fully consistent with the other results presented here.

In summary, the data now available point to two general conclusions. (1) In the absence of an actual change in the pattern of impulse activity reaching a muscle, cross reinnervation has no significant effect. (2) A change in the pattern of impulse activity unaccompanied by a change in the motor innervation can produce effects equalling or surpassing those brought about by nerve cross union.

Implications for nature of neural influence

These observations rule out the possibility that the physiological and biochemical differences between fast and slow muscles are attributable to fundamental differences in the chemotrophic character of their motor nerves. Therefore in seeking an explanation of the neural influence on muscle differentiation we may confine our attention to impulse-dependent mechanisms. The data are not incompatible with the secretion of a trophic substance or substances by those α motoneurones, irrespective of type, which sustain a particular pattern of activity[19]. Such an hypothesis, however, now seems unnecessarily elaborate, for we have shown that the same phenomena can be explained just as adequately in terms of events taking place wholly within the muscle. According to this interpretation, gene expression in skeletal muscle is subject to the regulatory influence of one or more substances whose intracellular concentrations alter as a consequence of changes in the level of contractile or metabolic activity. Such a shift in emphasis from explicit chemical control by the central nervous

system to adaptive regulation by the muscle may prove to be a more fruitful concept in the design of future experiments in this area.

1 Eccles, J. C., Eccles, R. M., and Lundberg, A., *J. Physiol., Lond.*, **142**, 275–291 (1958).
2 Buller, A. J., Eccles, J. C., and Eccles, R. M., *J. Physiol., Lond.*, **150**, 399–416 (1960).
3 Salmons, S., *J. Physiol., Lond.*, **188**, 13–14P (1967).
4 Salmons, S., thesis, Univ. Birmingham (1968); *Br. Sci. News*, No. 83, 2–4 (1971).
5 Salmons, S., and Vrbová, G., *J. Physiol., Lond.*, **201**, 535–549 (1969).
6 Sréter, F. A., Romanul, F. C. A., Salmons, S., and Gergely, J., in *Exploratory Concepts in Muscular Dystrophy*, **2** (edit. by Milhorat, A. T.), Int. Cong. Series No. 333, 338–343 (Excerpta Medica, Amsterdam, 1974).

7 Romanul, F. C. A., Sréter, F. A., Salmons, S., and Gergely, J., in *Exploratory Concepts in Muscular Dystrophy*, **2**, (edit. by Milhorat, A. T.), Int. Cong. Series No. 333, 344–348 (Excerpta Medica, Amsterdam, 1974).
8 Sréter, F. A., Elzinga, M., Salmons, S., Luff, A. R., and Mabuchi, K., *FEBS Lett.*, **57**, 107–11 (1975).
9 Buller, A. J., Eccles, J. C., and Eccles, R. M., *J. Physiol., Lond.*, **150**, 417–439 (1960).
10 Yellin, H., *Nature*, **249**, 596 (1974).
11 Close, R., in *Contractility of Muscle Cells and Related Processes* (edit. by Podolsky, R. J.), 175–188 (Prentice-Hall, London, 1971).
12 Weeds, A. G., Trentham, D. R., Kean, C. J. C., and Buller, A. J., *Nature*, **249**, 596 (1974).
13 Peckham, P. H., thesis, Case Western Reserve Univ. (1972); Peckham, P. H., Mortimer, J. T., and Van Der Meulen, J. P., *Abstr. third A. Mtg. biomed. Engng. Soc.*, Baltimore (1972).
14 Al-Amood, W. S., Buller, A. J., and Pope, R., *Nature*, **244**, 225–227 (1973).
15 Sréter, F. A., Gergely, J., Salmons, S., and Romanul, F. C. A., *Nature new Biol.*, **241**, 17–19 (1973).
16 Romanul, F. C. A., Sréter, F. A., and Salmons, S., *Trans. Am. neurol. Ass.*, **98**, 297–298 (1973).
17 Salmons, S., and Sréter, F. A., *J. Anat.*, **120**, 413 (1975).
18 Sréter, F. A., Luff, A. R., and Gergely, J., *J. gen. Physiol.*, **66**, 811–821 (1975).
19 Buller, A. J., and Lewis, D. M., *J. Physiol., Lond.*, **178**, 343–358 (1965).
20 Sréter, F. A., *Archs Biochem. Biophys.*, **134**, 25–33 (1969).
21 Sarkar, S., Sréter, F. A., and Gergely, J., *Proc. natn. Acad. Sci. U.S.A.*, **68**, 946–950 (1971).

letters to nature

Spectral characteristics of transient X-ray sources

THE transient X-ray sources A1524−62 and A0535+26 were briefly observed by the University College, London, proportional counter spectrometer, experiment C on the Ariel V satellite (P. W. Sanford and J. C. Ives, unpublished). The resulting spectra help to complete the spectral history of these two sources, making seven sources for which there is now extensive spectral information. A general scheme emerges from these in which the sources with hard spectra ($n \lesssim 2$) show little spectral evolution and those with soft spectra ($n > 2$) exhibit marked softening with time. We discuss here this subdivision in relation to other properties of the sources.

The X-ray transient A1524−62 was first reported by Pounds[1] and a detailed light curve has been presented by Kaluzienski *et al.*[2]. These authors reported that the spectrum before peak intensity showed a slight softening, with the data at the precursor maximum being well fitted by a power law spectrum of photon index −2.5 and with no measurable soft X-ray absorption ($< 10^{22}$ atoms cm^{-2}, 1σ). Their data were not consistent with an isothermal Bremsstrahlung X-ray spectrum.

This source was observed by experiment C for two days, February 6–8, 1975, well after peak intensity, and the mean intensity in the 3–6-keV energy band was found to be 0.183 ± 0.001 photons cm^{-2} s^{-1} (1σ), in good agreement with Kaluzienski *et al.*[2]. The orbit-by-orbit intensities showed no variability

outside the 3σ uncertainties, setting an upper limit of 10% on intensity changes at the time scale of hours.

The spectrum obtained by integrating the data over this period is shown in Fig. 1 together with the best fitting power law spectrum. The best fitting parameters for both power law and exponential trial spectra are given in Table 1. The power law spectrum (index 4.23 ± 0.07) gives a better fit to the data,

Fig. 1 The spectrum obtained during the 2-d observation of A1524−62. The continuous curve is the best fitting power law spectrum whose parameters are listed in Table 1.
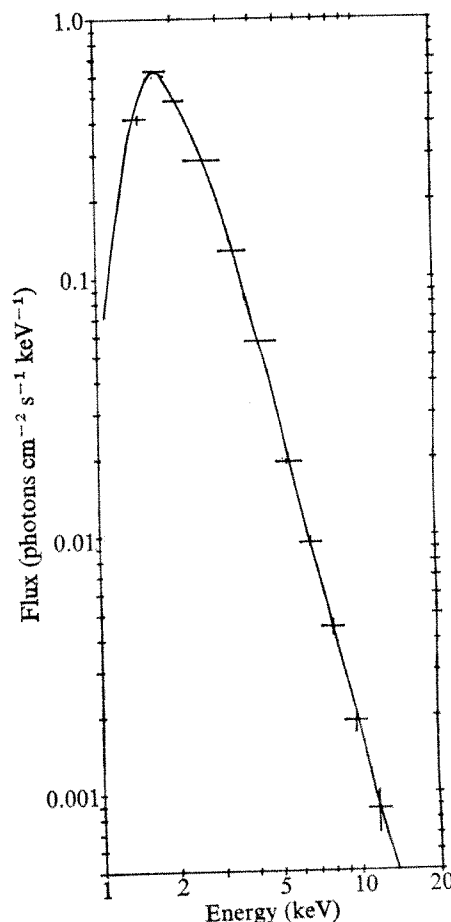
Table 1 Best fitting spectral parameters for A1524−62 and A0535+26

	Power law spectrum	Exponential spectrum
A1524−62 A0535+26	$dN/dE = AE^{-n}\exp(-N_H\sigma)$	$dN/dE = A\exp(-E/kT)$ $E^{-1}\exp(-N_H\sigma)$
A	30.6 ± 3.0	5.2 ± 0.5
n or kT	4.23 ± 0.07	1.43 ± 0.05
N_H	$(3.40 \pm 0.07) \times 10^{22}$	$(7.1 \pm 0.8) \times 10^{21}$
χ^2 for 8 d.f.	256	465
A	1.3 ± 0.08	Data not
n or kT	0.67 ± 0.03	consistent with
N_H	No detectable cutoff ($< 1 \times 10^{22}$)	single temperature fit
χ^2 for 29 d.f.	14	

σ is the Brown and Gould[3] absorption cross section.
The errors are 1 s.e., and are calculated by the least squares program used to fit the trial spectra.

Table 2 Summary of the characteristics of the seven most extensively observed transient X-ray sources

Source	τ_r (d)	τ_1 (d)	τ_2 (d)	Spectral parameters* Rise	Peak	Decay 1	Decay 2	I_{max} (erg cm^{-2} s^{-1})	D (kpc)	L_{max} (erg s^{-1})	Comments (OC= Optical candidate)
Cen X-4 refs 7, 8	4	6	40	n 1	2	3 (3.5 keV)	3 (3.5 keV)	5×10^{-7}			
3U1543−47 refs 9, 10	4	10	100	n 3 (3.5 keV) N_H	4 $< 10^{22}$	4	4 (2 keV) 6×10^{21}	4×10^{-8}	3	4×10^{37}	OC = Variable (ΔM = 0.06) M giant
A1524−62 see text	4	30	60	n 1.5–2.5 N_H	$< 10^{22}$		4.2 3.4×10^{22}	5×10^{-9}			
A1118−61 refs 11–13	3	7		n 1.0† N_H 4×10^{22}	0.9† 4×10^{22}	1.1† 4×10^{22}		4×10^{-9}	3	4×10^{36}	Pulsating (P = 405 s). OC = Variable (ΔM = 0.1) Be Star
A1743−28 refs 14, 15		18	90	n N_H		3 10^{23}	3 10^{23}	3×10^{-8}	10‡	3×10^{38}	
A0535+26 refs 16, 17	6	20		n 0.67 N_H $< 10^{22}$	0.8 $< 10^{22}$	1.1 $< 10^{22}$		5×10^{-8}	1	5×10^{36}	Pulsating (P = 104 s). OC = Variable (ΔM = 0.7) Be Star
A0620−00 refs 18, 19	5			n 1.6 N_H $< 10^{22}$	5 5×10^{22}	5 (1.3 keV) 3.9×10^{21}		1×10^{-6}	2	4×10^{18}	OC = Blue star, varied by $\Delta M \simeq 8$ in coincidence with X-ray outburst

τ_r = Time for intensity to increase from $(I_{max}\,e^{-1})$ to I_{max}: τ_1 = time for intensity to decrease by a factor e after maximum intensity: τ_2 = sources Cen X-4, 3U1543−47, A1524−62, A1742−28 exhibited a second slower decay of e-folding time τ_2. Sources Cen X-4, 3U1543−47 showed a further rapid decay, after $(1–2) \times \tau_2$, leading to their disappearance: I_{max} = intensity at maximum; D = estimated distance; L_{max} = luminosity at maximum.
*For comparison, all spectra are described in terms of photon indices. Where not quoted directly, it has been roughly calculated from the reported temperature, which is also quoted in parentheses.
†Unpublished revised values of n and N_H, following calibration of Ariel V experiment C on the Crab Nebula.
‡Distances have been estimated from proposed optical identifications, except for A1742−28, the galactic centre nova, for which the 10 kpc is based on the celestial position and observed high column density[2].

but both spectra require significant soft X-ray absorption, equivalent, in the case of the power law fit, to a neutral hydrogen column of $(3.40 \pm 0.07) \times 10^{22}$ atoms cm^{-2}. The spectrum thus shows that a considerable spectral softening occurred in the source at or after maximum intensity.

The transient A0535+26 was observed by one or more of the Ariel V satellite experiments during the period April 13–June 6, 1975, and again between June 21 and August 15, 1975. The light curve (A. P. Willmore, unpublished) is fairly typical of other transient X-ray sources, with a precursor peak, a rise to maximum intensity (on May 1), followed by a quasi-exponential decay (time constant ~20 d). A subsidiary peak of 15% of the maximum has also been reported.

A0535+26 was observed by experiment C on April 28, that is, before maximum. The resultant data are shown in Fig. 2, together with the higher energy data from the crystal scintillator experiment on Ariel V (ref. 4). The experiment C data are best fitted by a power law spectrum with photon index of 0.67 ± 0.03, without measurable cutoff (see Table 1). However, neither a power law spectrum nor a single temperature exponential spectrum is consistent with the combined low and high energy data. Ricker et al.[5], from a balloon experiment on June 1, also found spectral steepening at higher energies not consistent with a single temperature exponential spectrum.

Comparison of the spectral data between 2 and 20 keV obtained by experiment C just before maximum, and by Ricketts et al.[6], suggests a slight spectral softening during the decay at these energies, but not so marked as for other X-ray transients.

The information on the most extensively observed transients is summarised in Table 2. A detailed review concerning the 5 most recent sources has been given by Willmore (unpublished). We stress that the values reported in Table 2 not directly given in the quoted literature, but estimated on the basis of the available information, are only indicative. This applies to the majority of the quoted time scales and to some of the spectral indices.

All the time scales are broadly similar, ranging from a few weeks to a few months, but this probably arises from observational selection effects. In fact there are indications of both faster[20] and slower[21] transients. Nonetheless, we are probably discussing a group of substantially similar events. Among the theories proposed for 'X-ray novae', variable accretion on to

compact objects in binary systems (refs 10, 15, 20, 22, 23, and F. K. Li, G. F. Sprott and G. W. Clark, unpublished) seems the most natural, in view of the similarities with known binary X-ray sources and the independent evidence of intermittent mass exchange in binary systems. The time scales should then be associated with the mechanism that drives the mass transfer.

Fig. 2 The spectrum of A0535+26. The experiment C data points for day 118, 1975 (1.5–15 keV) are combined with those reported by Coe *et al.*[4] (20–120 keV) for days 116–118, 1975.
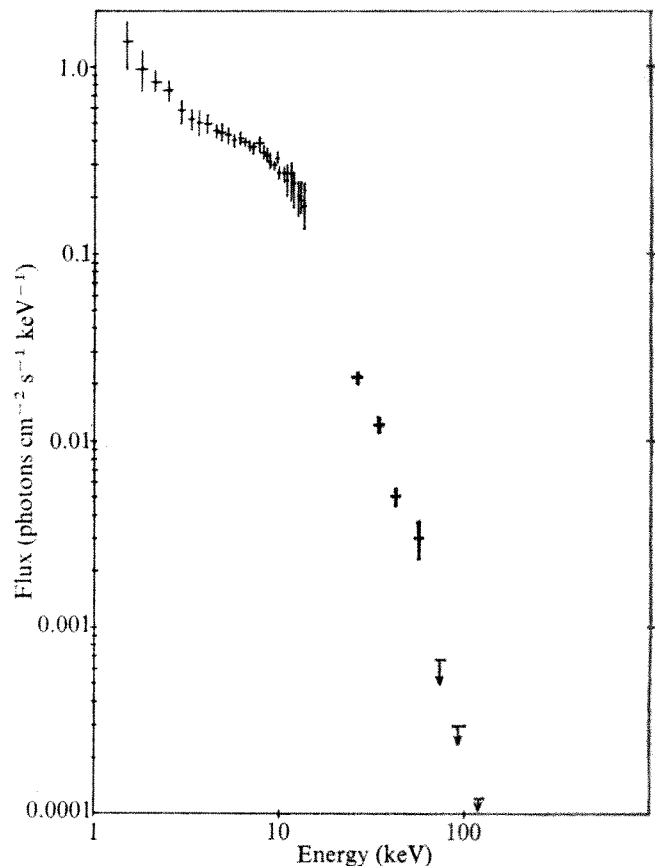
It is apparent from Table 2 that, in their spectral characteristics, A1118−61 and A0535+26 are similar to each other and different from the other five sources. The first two have very hard spectra ($n \simeq 1$) which do not show strong variation with phase. The other five have spectral indices at, or after, maximum of $n \simeq 3$–5 and in four cases where spectral observations before maximum are available, there is evidence that the spectrum has varied, evolving from a hard to a soft state. In three cases the variation is from $n \simeq 1$ to $n \simeq 3$–5.

Moreover, the two hard sources exhibit regular pulsation and the proposed optical counterparts are Be Stars ($M \gtrsim 10\ M\odot$), while for the other five pulsation has not been observed and the available optical identifications (for 3U1543−47 (ref.24) and A0621−00 (ref. 25)) indicate low mass companions. (The absence of further possible identifications in this group is consistent with this.)

It has been pointed out (L. Maraschi, A. Treves and E. P. J. van den Heuvel, unpublished) that among all galactic X-ray sources there exists a strong correlation between the occurrence of pulsation, the hardness of the X-ray spectrum and the association with systems containing bright early type stars. The analogous grouping within the more limited sample of transients therefore acquires significance, and can be interpreted simply as reflecting the difference between the two classes of systems, of which Cen X-3 and Sco X-1 can be taken as prototypes. It is not clear why the spectral type of the primary star should be related to the properties of the X-ray emission from the compact companion and a possible model is discussed by Maraschi *et al*. The different spectral evolution of transients of the two groups is an important fact to be considered in models accounting for the different X-ray characteristics of the the two types of binaries.

L. MARASCHI

Laboratorio di Fisica Cosmica e Tecnologie
Relative del CNR,
Istituto di Fisica dell'Università Milano,
Milano, Italy

H. E. HUCKLE
J. C. IVES
P. W. SANFORD

University College London,
Mullard Space Science Laboratory,
Holmbury St Mary,
Dorking, Surrey, UK

1 Pounds, K. A., *IAU Circ.*, No. 2729 (1974).
2 Kaluzienski, L. J., *et al.*, *Astrophys. J. Lett*, 201, L121 (1975).
3 Brown, R. L., and Gould, J., *Phys Rev. D.*, 1, 2252, (1970).
4 Coe, M. J., Carpenter, G. F., Engle, A. R., and Quenby, J. J., *Nature*, 256 630 (1975).
5 Ricker, G. R, *et al.*, *Astrophys. J. Lett.*, 204, L73 (1976).
6 Ricketts, M. J., Turner, M. J. L., Page, C. G, and Pounds, K A , *Nature*, 256, 631 (1975).
7 Evans, W. D., Belian, R. D , and Connor, J. P., *Astrophys. J. Lett.*, 159, L57 (1970).
8 Evans, W. D., Belian, R. D., and Connor, J. P., *Proc. Conf. Transient Cosmic and X-ray Sources, Los Alamos, New Mexico* (1973)
9 Belian, R. D., Connor, J. P., and Evans, W. D., *Proc. Conf. Transient Cosmic and X-ray Sources, Los Alamos, New Mexico* (1973)
10 Matilsky, T., Gacconi, R., Gursky, H., Kellogg, E., and Tananbaum, H., *Astrophys. J. Lett*, 174, L53 (1972).
11 Chevalier, C., and Ilovaisky, S. A., *IAU Circ.*, No. 2778 (1975).
12 Eyles, C. J., Skinner, G. K., Willmore, A. P., and Rosenberg, F. D., *Nature*, 254, 577 (1975).
13 Ives, J. C., Sanford, P. W., and Bell-Burnell, S. J., *Nature*, 254, 578 (1975).
14 Eyles, C. J., Skinner, G. K., Willmore, A. P., and Rosenberg, F. D., *Nature*, 257, 291 (1975).
15 Branduardi, G., Ives, J. C., Sanford, P. W., and Maraschi, L., *Mon. Not. R astr. Soc.*, 175, 47 (1976).
16 Liller, W., *IAU Circ.*, No. 2780 (1975).
17 Rosenberg, F. D., Eyles, C. J., Skinner, G. K., and Willmore, A. P., *Nature*, 256, 628 (1975).
18 Elvis, M., Page, C. G., Pounds, K. S., Ricketts, M. J., and Turner, M. J. L., *Nature*, 257, 656 (1975).
19 Ricketts, M. J., Pounds, K. A., and Turner, M. J. L., *Nature*, 257, 657 (1975).
20 Kellogg, E., Gursky, H., Murray, S., Tananbaum, H , and Gacconi, R., *Astrophys J. Lett* , 169, L99 (1971).
21 Clark, G. W., *Lectures at the Enrico Fermi Summer School on Physics and Astrophysics of Neutron Stars and Black Holes*, (Varenna, Italy, 1975).
22 Fabian, A. D., *Mon. Not. R astr. Soc.*, 173, 161 (1975).
23 Paolini, F., and Shapiro, S. L., *Nature*, 255, 618 (1975).
24 Forman, W., and Liller, W., *Astrophys. J. Lett.*, 183, L117 (1973).
25 Ward, M. J., Penston, M. V., Murray, C. A., and Clements, E. D., *Nature*, 257, 659 (1975).

The nature of association of equatorial spread *F* with magnetic activity

STATISTICAL studies based on a published data of f_0F_s have revealed that equatorial spread *F* depends inversely on magnetic activity during the sunspot maximum period but tends to disappear or even become positive in some seasons during the sunspot minimum period[1-3]. Somayajulu *et al*.[4], using a different technique, considered the occurrence of spread *F* during different phases of magnetic storms, and they concluded that the initial storm phase does not have any inhibiting effect on spread *F* occurrence and that spread *F* occurs in general during the main phase of the magnetic storms. They observed that, during the recovery phase when the field value is quite different from the normal value, there is no spread *F* but it is present when the magnetic field value is close to its normal value. To investigate this discrepancy with the earlier results further, the spread *F* behaviour has been studied[4] in relation to K_p indices, and for the Trivandrum data of 1970, and though the overall correlation between spread *F* and magnetic activity is negative, occurrence of spread *F* was found to decrease initially with increasing ΣK_p and then to start increasing at higher ΣK_p values. The effect of magnetic activity on equatorial spread *F* is reported here.

Equatorial spread *F* starts generally around 1900 LMT and lasts for a few hours. On some occasions, however, it persists throughout the night[5]. We have examined quarter hourly ionograms, during onset and development phases of spread *F*, from Trivandrum (dip−0.6°S) for 1970, and from Kodaikanal (dip−3.5°N) for 1969 and 1964, and the *f*-plots of Huancayo (dip−1°N) for 1968 to study the characteristics of spread *F* during its onset and development phase with magnetic activity as indicated by K_p indices. For this, the K_p indices for 1800–2100 LMT of Trivandrum and Kodaikanal and 1900–2200 LMT of Huancayo were used. The shift of 1h for Huancayo was used because K_p as given in the Solar Geophysical Data corresponds to that particular 3-h interval. The daily sum of K_p indices, represented by ΣK_p, was also calculated. The spread *F* data were grouped according to K_p and ΣK_p and percentage spread *F* was estimated for each (Figs 1–3).

In Figs 1–3, vertical bars indicate the 95% confidence limits. The yearly average sunspot numbers for the year 1968, 1969 and 1970 are 106, 106 and 105 respectively. These three years correspond to the period around sunspot maximum in that particular solar cycle and the year 1964 corresponds to the sunspot minimum (with a sunspot number of 10). In the following discussion no distinction is made between 1968, 1969 and 1970 in regard to the sunspot activity.

The most striking features of Figs 1, 2 and 3 is the reversal of the trend of variation of occurrence of spread *F* with increasing K_p value. This starts decreasing with increasing K_p but at a particular value of K_p it starts increasing again. It is

Fig. 1 Percentage occurrence of spread *F* at Trivandrum (75°E) against (*a*) the K_p index during 1800–2100 LMT (correlation coefficient = −0.19) and (*b*) against ΣK_p for the whole day (correlation coefficient = −0.29).
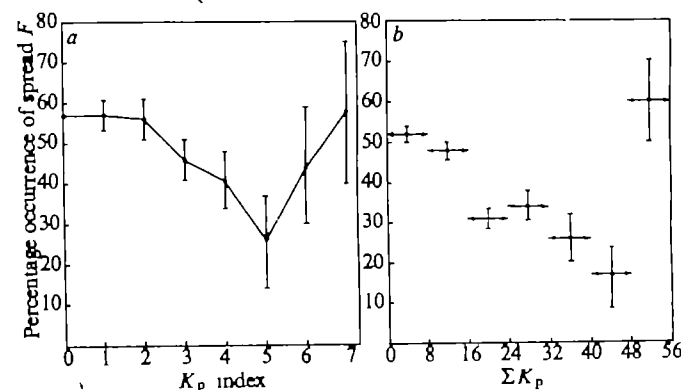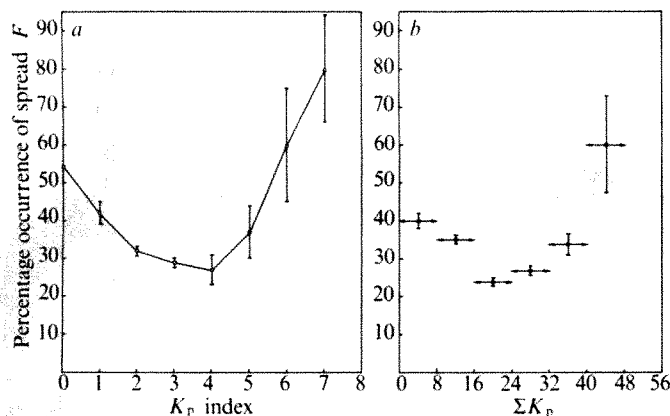
Fig. 2 Percentage occurrence of spread F at Huancayo (75°W) against (*a*) K_p index during 1900–2200 LMT (correlation coefficient $= -0.12$) and (*b*) ΣK_p during the whole day (correlation coefficient $= -0.1$).

even greater at the highest K_p value compared to that at the quietest condition ($K_p = 0$), as can be seen clearly from Figs 2 and 3. In Fig. 1 the two are almost equal. The value of K_p at which the reversal occurs at Huancayo is less than that at Trivandrum and Kodaikanal. Another important feature is that the decrease in spread F with K_p is rather slow initially at Trivandrum and Kodaikanal compared to that at Huancayo but later the opposite is true. For 1964 the general level of spread F is low at Kodaikanal as can be seen from Fig. 3. The reversal of the trend is seen as for the sunspot maximum year but with a major difference: it occurs at a much lower value of K_p($= 1$). The correlation coefficients between spread F and K_p index are also shown in Figs 1, 2 and 3. They are all negative except for 1964 at Kodaikanal. The coefficients are all $> 95\%$ confidence limit. This shows that though the relationship between K_p and spread F reverses at high K_p, the overall correlation between the two remains negative in accord with the earlier results[1-3]. The overall negative correlation obviously arises because days with high K_p value (or ΣK_p) are less numerous than those with low K_p value (or ΣK_p).

It is known that the onset of spread F is earlier in the period of sunspot maximum compared with its onset in sunspot minimum period. We have here considered the same 3-h

Fig. 3 Percentage occurrence of spread F at Kodaikanal (75°E) against the K_p index during 1800–2100 LMT. ○, Data for 1969, correlation coefficient $= -0.09$; □, data for 1964, correlation coefficient $= +0.12$.
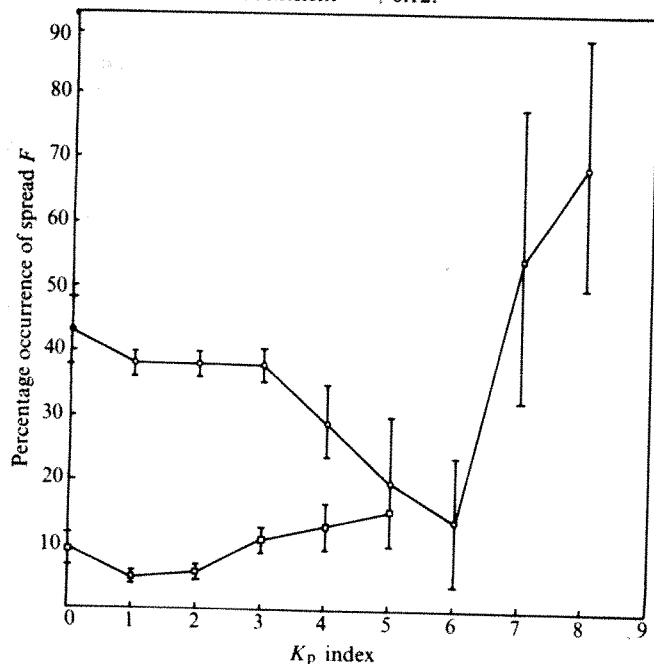
period (1800–2100 LMT) both for sunspot maximum and minimum years for Kodaikanal As such, the occurrence of spread F during the onset and development phase of spread F in the sunspot minimum period may be underestimated. Its variation with K_p may not, however, be affected.

The reversal at a lower K_p value for Huancayo during the sunspot maximum years is quite likely to be true for the sunspot minimum period also. One would then expect that the occurrence of spread F starts increasing right from $K_p = 0$. This may result in comparatively more spread F at Huancayo during sunspot minimum relative to that at sunspot maximum because there would be more days with low K_p than with high. In fact many investigators[3] reported just this.

V. V. SOMAYAJULU
B. V. KRISHNAMURTHY

Space Physics Division,
Vikram Sarabhai Space Centre, Trivandrum-695022, India

1 Huang, Y. N., and Yeh, C. Y., *J. atmos. terr. Phys.*, **32**, 1765 (1970).
2 Chandra, H., and Rastogi, R. G., *Ann. Geophys.*, **28**, 37 (1972).
3 Chandra, H., and Rastogi, R. G., *Ann. Geophys.*, **28**, 709 (1972).
4 Somayajulu, V. V., Sen Gupta, K., and Krishnamurthy, B. V., *Symp. Solar Planet. Physics, Phys. Res. Lab., Ahmedabad, India* (1976).
5 Somayajulu, V. V., Sen Gupta, K., and Krishnamurthy, B. V., *Nature*, **257**, 112 (1975).

v.l.f. emission from ring-current electrons

v.l.f. emissions associated with the enhancement of ring-current electrons during magnetic storms and substorms have been detected by the satellite S^3-A (Explorer 45) on an equatorial orbit[1,2]. The emissions observed near the geomagnetic equator consist of essentially two frequency regimes; one above the electron gyrofrequency at the equator, f_H—the electrostatic mode[3-5], which peak near $(n+\frac{1}{2})f_H$, where n is positive integer (emissions up to $n = 10$ have been observed[3])—and the other below f_H—the whistler mode, which has a conspicuous gap at exactly $f_H/2$. Here we describe the characteristics of this v.l.f. emission and the associated enhancements in electron intensity as well as the anisotropies of the ring-current electron distribution, and we also give an interpretation of the bimodal frequency distribution of the equatorial whistler mode emissions. This distribution has been known for the past few years but has not been well explained[6,7].

Figure 1 is a part of the wide-band data obtained during the main phase of the December 17, 1971 magnetic storm[1]. Figure 1*a* is the a.c. magnetic field data measured by the search-coil magnetometer with the upper cutoff near 3 kHz (ref. 8) and Fig. 1*b* is the a.c. electric field data obtained by the electric field sensor with the upper cutoff of 10 kHz (ref. 9). The figure (*a* and *b*) shows the time sequence of the observed emissions along the inbound orbit (No.101) of the satellite as f_H changes from ~ 3.0 kHz at 2000 UT to 6.0 kHz at 2100 UT. Although the automatic gain control of the sensor produces occasional spurious intensity variations, it should be noticed that emissions at frequencies $> f_H/2$ are generally stronger and more continuous than those $< f_H/2$.

Another example of the equatorial magnetospheric v.l.f. emissions is shown in Fig. 2, where the emission data, which were observed during the substorm on January 22, 1972, are digitised in 1-min bins and are normalised by the value of f_H computed from onboard magnetometer data. Shown in the upper portion of Fig. 2 is the magnetic latitude of the satellite at the time of the observations. The dark and light shades correspond to the electrostatic mode and the whistler mode emissions, respectively. The frequency of occurrence graph, on the right, shows that the whistler mode has a bimodal frequency distribution with a sharp dip at exactly $f_H/2$.

The variations of the differential electron intensities ($cm^{-2} s^{-1} sr^{-1} keV^{-1}$) for the same date are shown in Fig. 3 for electrons of 90° pitch angle and for electrons with the smallest observable pitch angle (close to 25°). The difference between these two plots for each energy is a manifestation of the pitch-angle anisotropy and is indicated by the vertical hatched lines.

Magnetic (Minnesota)

Electric (Iowa)

2000 UT DAY 351
$L = 5.044$
M Lat $= +1.7°$
MLT $= 21.6$

2030 UT
$L = 4.739$
M Lat $= +1.4°$
MLT $= 22.0$

2100 UT
$L = 4.288$
M Lat $= +0.9°$
MLT $= 22.6$
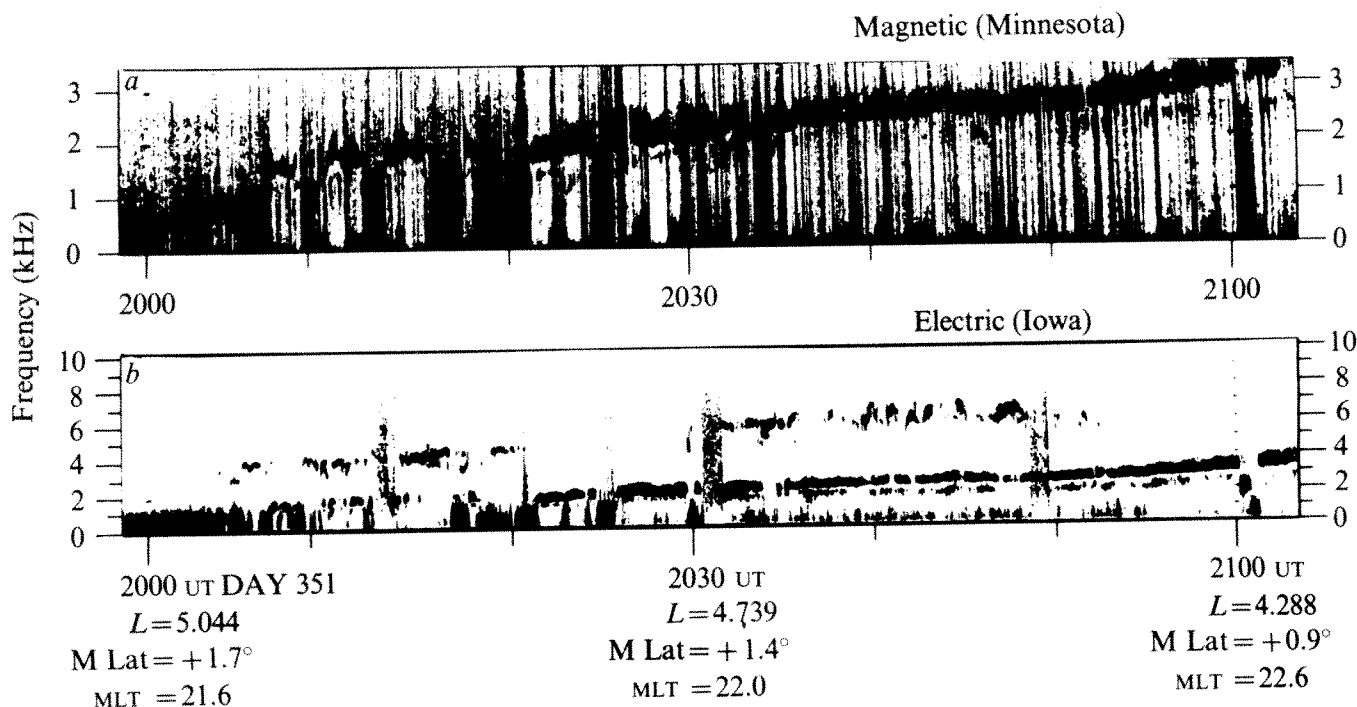
Fig. 1 A part of the v.l.f. emission event during the main phase of the geomagnetic storm on December 17, 1971. The emission of an a.c. magnetic field ($f \leqslant 3$ kHz) and an a.c. electric field ($f \leqslant 10$ kHz), are indicated in (*a*) and (*b*) respectively. The numbers shown at the bottom are: (1) the universal time, UT; (2) a measure of geocentric distance of the satellite in the equatorial plane, L (in R_E units); (3) the magnetic latitude, M Lat; (4) the magnetic local time, MLT. The electron gyrofrequency, f_H, changes from 3,067 kHz at 2000 UT to 3,662 kHz at 2030 UT and to 5,986 at 2100 UT.
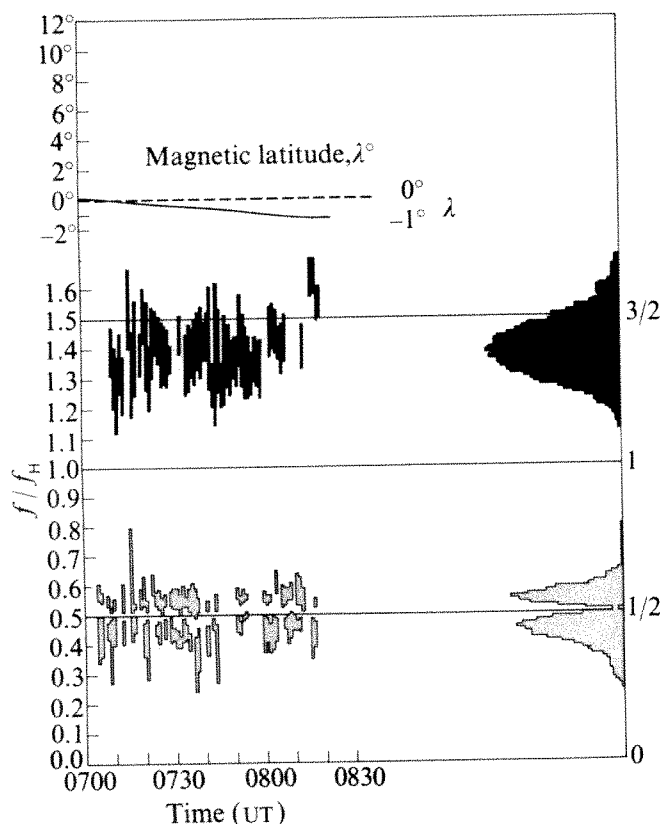
Fig. 2 The frequency distribution of the v.l.f. emissions observed by the S³-A satellite during the substorm on January 22, 1972. The frequency distributions made by accumulating the number of occurrence in each frequency interval ($\Delta f/f_H = 0.01$ is used) are shown on the right side of the figure. There is a conspicuous dip in the bimodal frequency distribution of the whistler-mode emission at exactly $f_H/2$.

The times indicated by horizontal hatched lines and grey shade are the periods when the emissions shown in Fig. 2 were observed and when the satellite is inside the plasmasphere, respectively. A short period around 0700 UT (approximately 2000 MLT—inside the plasmasphere) represents the double structure of the evening side plasmapause[2].

We can see that the v.l.f. emissions are observed only when the satellite is outside the plasmasphere and that the beginning of the emission coincides with the satellite's encounter with the large electron fluxes there. It should be noticed, however, that the increase in the electron intensities associated with this observed v.l.f. emission is limited to the low energy electrons only (< 6 keV for this event).

To see the energy spectrum of electrons, the flux spectrum at 0730 UT on January 22, 1972 is plotted in Fig. 4. The mean quiet period spectrum observed by this satellite, which has been given by Lyons and Williams[10] is also plotted for comparison. Approximating these energy spectra and the pitch angle distribution of electrons by

$$j(E, \alpha) \sim E^{-n} \sin^m \alpha \qquad (1)$$

The parameters n and m for three energy regimes are also shown in Fig. 4. We can see that not only the intensity increase but also the anisotrophy is larger in the lower energy electrons.

The growth rate of waves excited by resonant electrons increases with increasing anisotropy of the electrons, and the peak of emission shift to higher frequency as the anisotropy of the resonant electrons increases, particularly for the soft spectrum electrons. For electrons obeying equation (1), the peak of emission is $> f_H/2$ for $m \gtrsim 2$ with $n \gtrsim 1$ (ref. 11). The emission above $f_H/2$ can be also excited by the cyclotron instability of a plasma composed of thermal and quasithermal (20–50 eV) particles if the anisotropy of the quasithermal plasma is suitably large[12]. The S³-A satellite did not, however, carry detectors for these low energy plasma particles.

Since the resonance energy, E_R, for electrons encountering the

whistler waves of frequency, f, is given by

$$E_R = E_B \frac{f_H}{f}\left(1 - \frac{f}{f_H}\right)^3 \qquad (2)$$

where

$$E_B = \frac{B^2}{8\pi N} \quad \text{or} \quad E_B = 250\left(\frac{f_H}{f_p}\right)^2 \text{ (keV)} \qquad (3)$$

B, N and f_p are the magnetic field intensity, plasma density (cm^{-3}) and plasma frequency (kHz), respectively, the energies of electrons resonating with the whistler waves of frequency $f > f_H/2$ are always smaller than those for $f < f_H/2$. In this sense, the large electron intensities shown in Figs 3 and 4 are able to excite only the emissions of higher frequencies, $f \gtrsim f_H/2$.

Fig. 3 Variations of the directional differential electron intensities (cm^{-2} s^{-1} sr^{-1} keV^{-1}) on January 22, 1972. The two lines for each energy correspond to the intensities at pitch angle 90° ($\times$) and of the smallest observed pitch angle (close to 25°) (●).
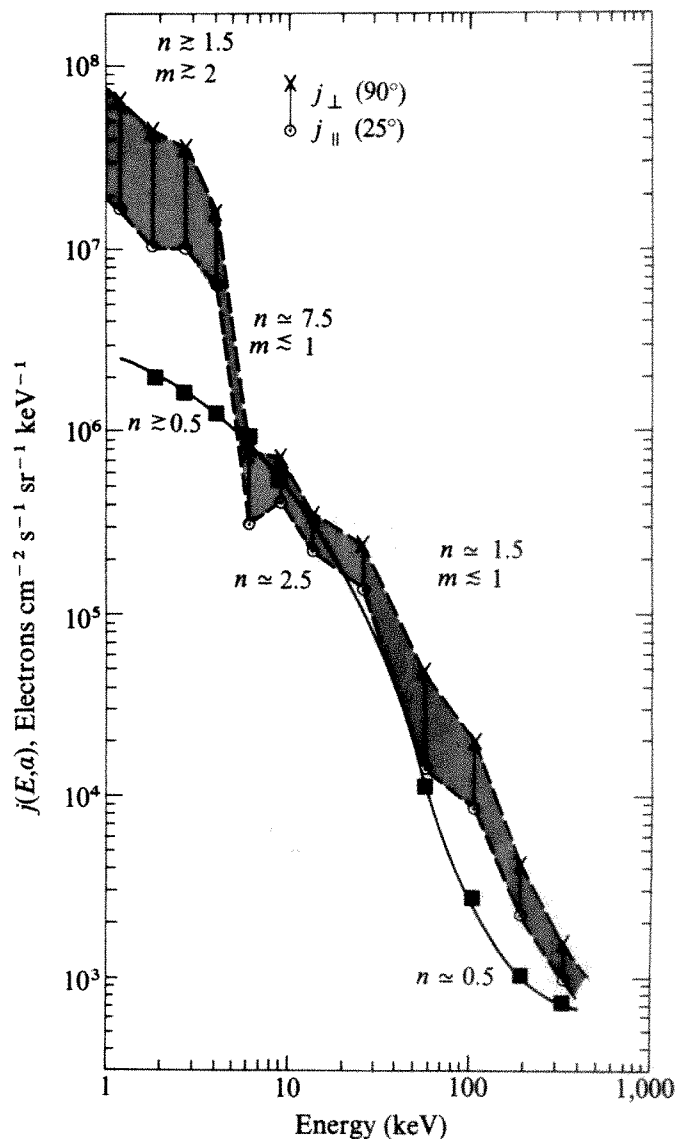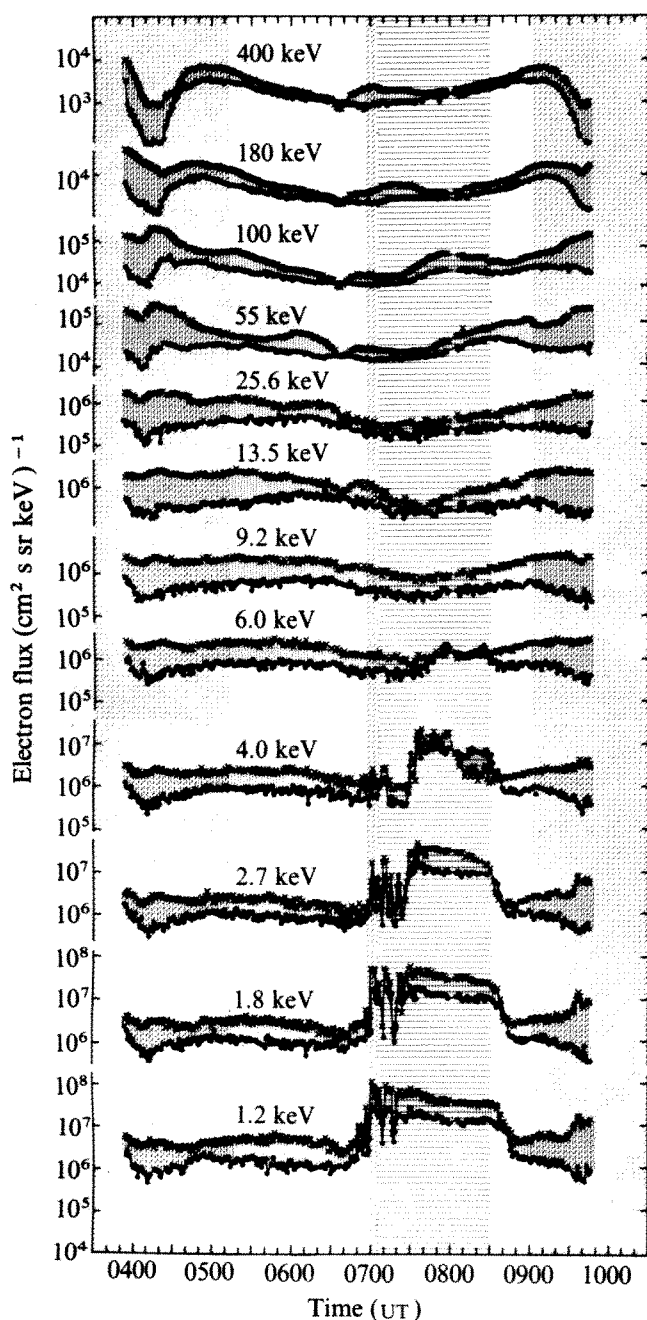
Fig. 4 Differential energy spectra of the enhanced ring current electrons during the v.l.f. emission event compared with the mean quiet time spectrum (■)[10]. The shaded part between two directional intensities ($\times$ for 90°, ○ for 25°) shows the anisotropy of the increased electron intensities at different energies. The parameters n and m for three energy regimes are also indicated in the respective parts of the curves.

On the other hand, because of the ~11° tilt of the Earth's magnetic dipole axis from the rotation axis and the small inclination (~4°) of the S³-A satellite orbit, emission events are occasionally observed off the equator. Figure 5 shows one example which was observed during the substorm on January 18, 1972. The notation is the same as in Fig. 2, and as indicated by the line in the upper left, the magnetic latitude of the satellite during the event was >10°.

From this figure, we can see that the frequency distribution >$f_H/2$ differs significantly from those of the near equatorial events, indicating that the off-equator propagation of this >$f_H/2$ magnetospheric emission is non-ducted. On the other hand, the emission <$f_H/2$ is similar to that observed near the equator. It seems therefore that these waves are duct-propagated with an upper cutoff of $f_H/2$, which was known from the investigation of whistler propagation in both ground and satellite observations[13-15].

Based on the information described above, the formation of the bimodal frequency distribution of the equatorial whistler mode emissions with a sharp dip at $f_H/2$ can be explained as follows: (1) The emission >$f_H/2$ is produced near the equator by
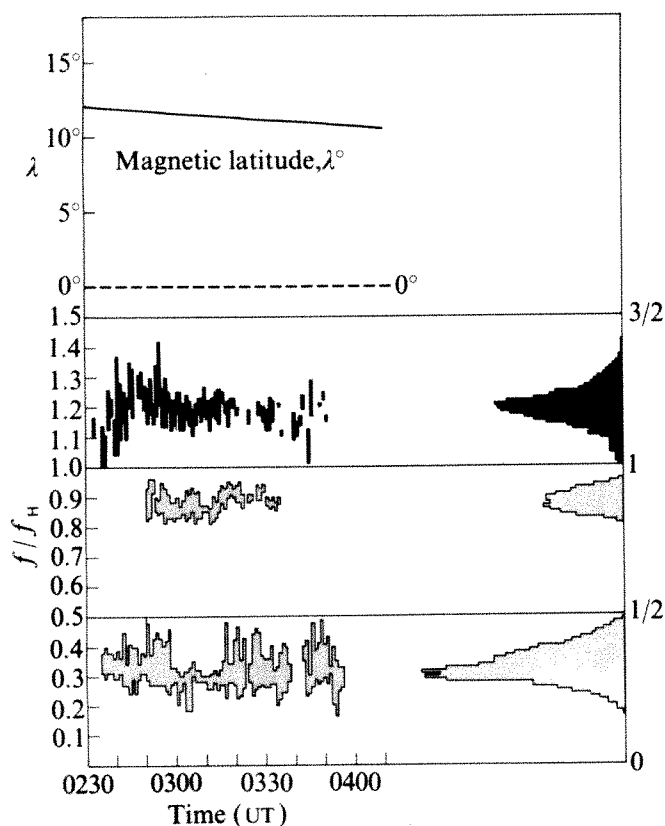
Fig. 5 The normalised frequency distribution. In this off-equatorial emission event, the emissions $> f_H/2$ are weaker than the emissions $< f_H/2$ and the frequencies are quite different from the near equatorial results shown in Figs 1 and 2.

the enhanced ring-current electrons, resonating with the weak broad band whistler-mode noises which are found quite generally in the region of $3 \lesssim L \lesssim 10$ of the Earth's magnetosphere; (2) the emissions $< f_H/2$ are produced in a similar manner near the equator but slightly outside the observing location, where the magnetic field is weaker, the corresponding gyrofrequency, f_H', is significantly smaller than that at the observing location, f_H; (3) those locally produced emissions with $f > f_H'/2$ propagate unducted off the equator in the magnetosphere, and are reflected back towards the equator from the regions where the lower hybrid resonance frequency, f_{LHR}, exceeds the frequencies of

these waves[16], as depicted schematically in Fig. 6; (4) the reflected waves propagate toward the equator in the narrow density-enhanced ducts whose upper cutoff is $f_H/2$ (refs 13–15).

Since $f_H' < f_H$, we can find, for example, that the emission produced at a frequency $f \simeq 0.55 f_H'$ (that is, $f > f_H'/2$) arrives at the observing location with $f = 0.45 f_H$ (that is, $f < f_H/2$), as indicated in Fig. 6.

This idea for the formation of the bimodal frequency distribution with a dip at $f_H/2$ (which is also called "the band of missing emissions" in geomagnetic equatorial whistler emissions, such as chorus[1,2,6]) has other support. Data gathered by the OGO-3 satellite when crossing the geomagnetic equator on January 21, 1967 (ref. 17), are shown in Fig. 7. The magnetic latitudes corresponding to Fig. 7a, b, c are $\sim 0°$, $-2°$ and $-5°$, and the arrow on the left side of each indicates the frequency corresponding to $f_H/2$. From these figures, we can see that: (1) emission at $f > f_H/2$ is stronger and more continuous than at $f < f_H/2$, indicating that the emission at $f > f_H/2$ is observed rather close to its source; (2) emission at $f < f_H/2$ is striated, indicating that this emission propagates in narrow ducts of enhanced ionisation; (3) there is no coherence between the upper and lower frequencies, except for the bottom example, (which is rather non-equatorial), indicating that the sources of those two are different, or at least separated.

Further aspects of our theory, including the growth rate of the emissions based on the observed large low energy electron intensity and anisotropy, and the ray tracing of non-ducting emissions produced near the equator with frequencies $f > f_H/2$, are being studied. Finally, there remain effects which are not well understood in the formation of the missing emission band carried by the interference of two modes of propagation and by the Landau damping for a frequency $\sim f \lesssim f_H/2$ (refs 6, 11 and 17). Investigations on the relation between the whistler-mode emission and the electrostatic-mode emission might produce more information about these problems.

KAICHI MAEDA
PAUL H. SMITH

NASA–Goddard Space Flight Center,
Greenbelt, Maryland 20771

ROGER R. ANDERSON

Department of Physics and Astronomy,
University of Iowa,
Iowa City, Iowa 52242

1 Maeda, K., *Planet. Space Sci.*, **24**, 341 (1976).
2 Anderson, R. R., and Maeda, K., *J. geophys. Res.* (in the press).
3 Kennel, C. F., Scarf, F. L., Fredericks, R. W., McGehee, J. H., and Coroniti, F. V., *J. geophys. Res.*, **75**, 6136 (1970).
4 Bernstein, I. B., *Phys. Rev.*, **109**, 10 (1958).
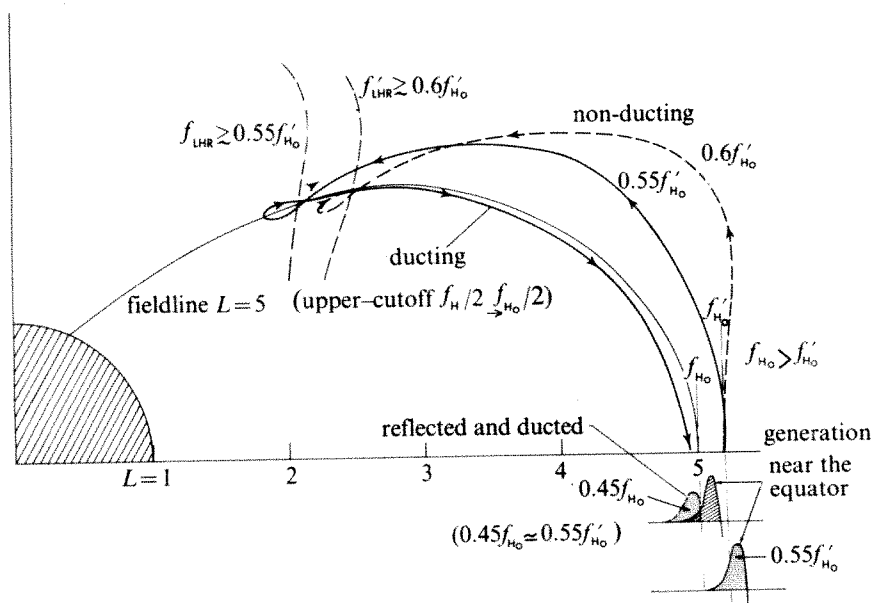
Fig. 6 A schematic depiction of the formation of the bimodal frequency distribution in the magnetospheric v.l.f. emissions observed near the equator. (All ray paths are outside the plasmasphere and the loops of the propagation path at reflection points are exaggerated.)
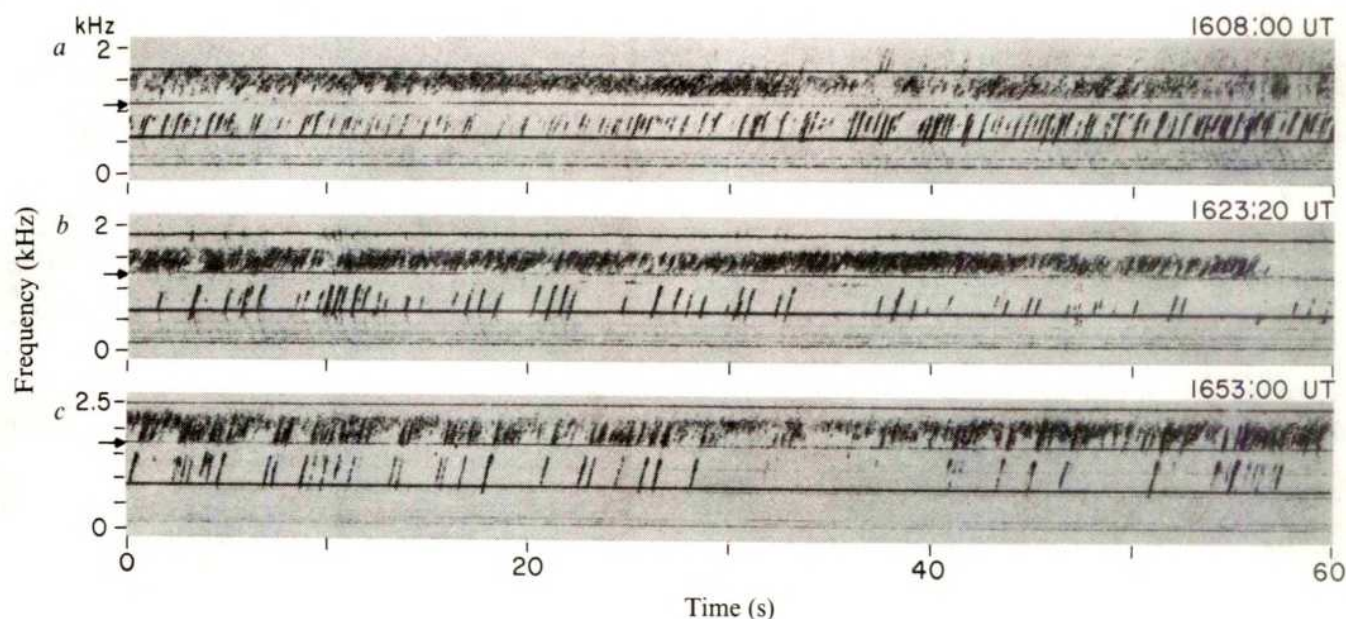
[5] Stix, T. H., *Theory of Plasma Waves* (McGraw Hill, New York, 1962).
[6] Tsurutani, B. T., and Smith, E. T., *J. geophys. Res.*, 79, 118 (1974).
[7] Gendrin, R., *Space Sci. Rev.*, 18, 145 (1975).
[8] Parady, B., and Cahill, L. J., Jr, *J. geophys. Res.*, 78, 4765 (1973).
[9] Anderson, R. R., and Gurnett, D. A., *J. geophys. Res.*, 78, 4756 (1973).
[10] Lyons, L. R., and Williams, D. J., *J. geophys. Res.*, 80, 943 (1975).
[11] Liemohn, H. B., *J. geophys. Res.*, 72, 39 (1967).
[12] Matsumoto, H., and Kimura, I., *Planet. Space Sci.*, 19, 567 (1971).
[13] Smith, R. L., Helliwell, R. A., and Yabroff, I., *J. geophys. Res.*, 65, 815 (1960).
[14] Angerami, J. J., *J. geophys. Res.*, 75, 6115 (1970).
[15] Helliwell, R. A., *Whistlers and Related Ionospheric Phenomenon* (Stanford University Press, 1965).
[16] Kimura, I., *Radiosci.*, 1, 269 (1966).
[17] Burtis, J. W., *Tech. Rep.*, 3469-3 (Stanford Radioscience Laboratories, 1974).

Gas movement through sea ice

UNLIKE ices from pure or freshwater, sea ice is a highly permeable medium for gases. We have found that: (1) the migration of these gases along the grain boundaries (probably in brine channels) was 2 to 6 times as great as that at right angles to the principal axis of the grain boundaries; (2) the rate of penetration is ~ 30 cm h⁻¹ at −15 °C for halogenated gases (in nl ml⁻¹ quantities) and 60 cm h⁻¹ at −7 °C for carbon dioxide (in μl ml⁻¹ quantities); (3) the vertical migration is about twice as fast as horizontal migration at −15 °C, and (4) for one experiment with a block of semi-fresh pressure ridge ice, the migration rate was ~ 60 cm h⁻¹ at −15 °C. The permeation constants are estimated to be 10^{-7} cm² s⁻¹ atm⁻¹ at −15 °C for SF_6, and 2×10^{-5} cm² s⁻¹ atm⁻¹ for CO_2 at −7 °C. These data strongly indicate that gas migration through sea ice is an important factor in ocean–atmosphere winter communication particularly when the surface temperature is > −10 °C.

It has been correctly reported[1] that pure water and lake water ice are effective barriers against any significant migration of gases at temperatures < −0.5 °C. Their permeabilities are 10^6 less than that of carbon dioxide through water, with lake water ice an order of magnitude more permeable because of traces of impurities in the ice. This is understandable when one considers results[2,3], which show that the size of veins at three grain intersections open or close drastically within a few hundredths to a few tenths of a degree below the freezing point for freshwater ices.

On the other hand, large brine loads would extend the temperature range over which these vein and brine channels would remain open to several degrees below the freezing point. The sea-ice temperature, of course, will be a gradient from −2 °C at the water–ice interface to that of the prevailing atmosphere at the air–ice interface.

We also observed brine to enter bore holes in noticeable quantities after 12 h at −15 °C: at −10 °C it was only slightly faster. At −6.5 °C brine temperatures, however, bore holes filled to the hydrostatic level within an hour. The bore holes were 50–75 cm deep in 2-m thick ice. Ink in brine was observed to travel 30 cm vertically within 30 min through a freshly quarried block of sea ice at −7 °C. There are also results[4,5] on the flow of water through ice at temperatures close to 0 °C.

Tracer gases (sulphur hexafluoride, perfluoromethyl bromide, and perfluoroethyl bromide) were resolved and detected in an electron capture gas chromatograph with a short Porpak N column. Carbon dioxide was analysed with a thermal conductivity device. Experiments with shorefast annual ice using halogenated tracer gases were carried out in April at the Naval Arctic Research Laboratory (NARL) at Barrow, Alaska, and other *in situ* experiments in pack ice were performed at the Arctic Ice Dynamic Joint Experiment (AIDJEX) Camp Caribou located ~240 miles NE of Barrow. Both sets of experiments provided the same rate values at −15 °C. Another set of *in situ* experiments at NARL in May, when the temperature was −7 °C, provided the data for CO_2. In early experiments, bore holes were charged with one or more of the halogenated tracer gases, and then the top of the hole was covered with a metal plate which was sealed to the surface with fresh water and snow. At intervals, cores were taken beside the sealed holes, leaving 25–30-cm ice walls. The cores were taken in the plane of the predominant ice grain boundaries from the gas-spiked holes, and perpendicular to it. These cores were immediately placed in gas-tight containers made from glass sewer pipe, and then flushed with He by means of stopcock fittings on the ends of the containers. In the laboratory, large loads of tracer gases were found in the head space gas from cores taken ~1.5 h after the start of the experiments. For all other experiments, 5-cm diameter bore holes were drilled to a depth of up to 1 m in the ice outside the work hut. Ice walls of 30–90 cm were left between three vertical holes. Metal plates with swagelock fittings were sealed to the ice over these holes with snow and freshwater. The two outside holes were manifolded through a closed

loop pump and switching system so that samples could be valved directly into the gas chromatograph. In other experiments, diagonal holes were bored in the ice so that 10- and 20-cm ice walls had to be penetrated vertically by the gases.

This work was supported by a grant from the Office of Naval Research.

THOMAS A. GOSINK
JOHN G. PEARSON
JOHN J. KELLEY

Institute of Marine Science,
University of Alaska,
Fairbanks, Alaska 99701

1 Hemmingsen, E., *Tellus*, 11, 355–359 (1959).
2 Langham, E. J., *Can. J. Earth Sci.*, 11, 1280–1287 (1974).
3 Raymond, C. P., and Harrison, W. D., *J. Glaciol.*, 14, 213–233 (1975).
4 Eide, L. I., and Martin, S., *J. Glaciol.*, 14, 137–154 (1975).
5 Wakahama, G., *et al.*, *Low Temp. Sci. A*, 31, 209–220 (1973).

On the marine geochemistry of cadmium

STUDIES of the geochemistry of trace metals in sea water are of great potential importance, given the wide diversity in their properties. Knowledge of their distribution provides information not only about the specific chemical processes in which they are involved but also about the formation, destruction and diagenesis of the major solid phases in which they become incorporated by biological and authigenic processes. Previous efforts to elucidate the behaviour of cadmium in the ocean have produced no systematic correlations with other oceanographic properties[1]. At the low levels encountered (< 0.1 parts per 10^9), problems with analysis and with contamination during sampling can be severe. Recent advances in analytical and handling techniques have led to progress in our understanding of other trace metals[2-4], hence a re-examination of the marine geochemistry of cadmium is now appropriate. This note describes the general distribution of cadmium in the ocean, based on three detailed vertical profiles from the Pacific.

Fig. 1 Cadmium, phosphate, and silicate profiles for three stations in the Pacific Ocean. ○, Silicate; ●, phosphate for (*a*) Stn 219 (*b*) Stn 226 and (*c*) Stn (293).
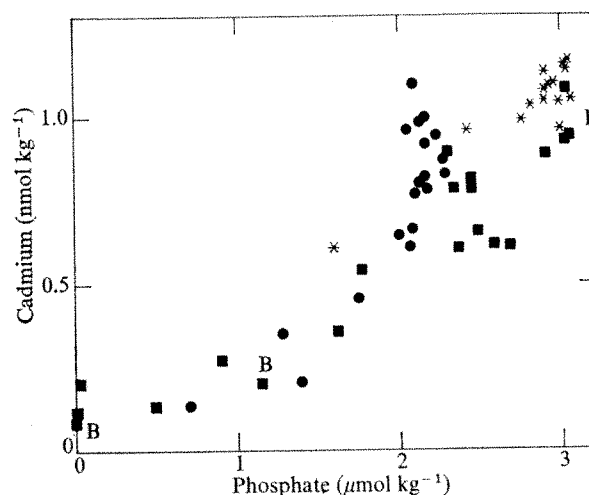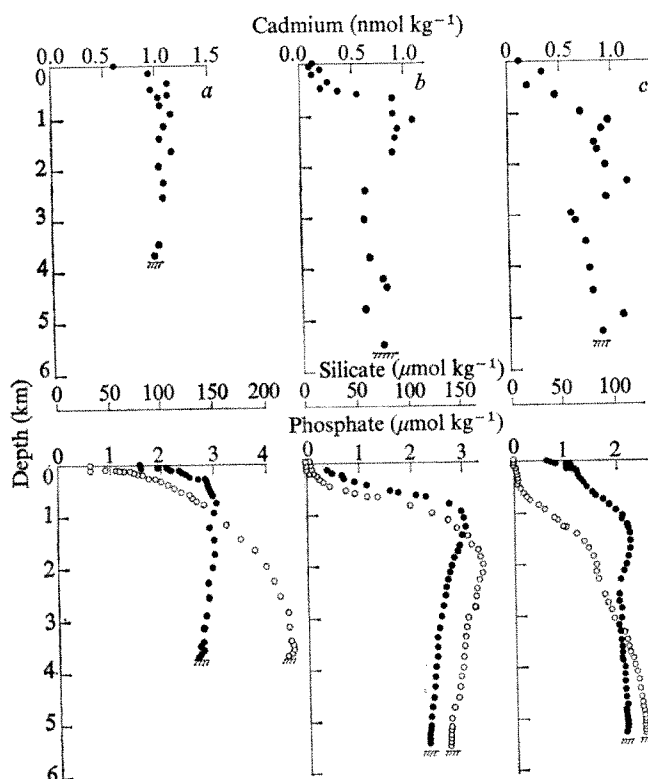
Fig. 2 Cadmium concentration against phosphate for three Pacific profiles (●, Stn 293; ✳, Stn 219; ■, Stn 226) and three data points from the literature, B (ref. 4) and K (ref. 6).

Station locations were chosen to cover a broad range of oceanographic conditions, from the south-western Pacific, S of New Zealand (Stn 293) to the temperate mid-Pacific, E of Japan (Stn 226) and the North Pacific upwelling region in the Bering Sea (Stn 219) (Table 1). The samples were taken using rosette-mounted 30-l PVC Niskin bottles with epoxy-coated internal springs and stored in acid-leached polyethylene containers, after acidification to $pH = 2$. The handling of containers and samples is described in detail elsewhere[3]. Cadmium was preconcentrated from 100-g samples by cobalt pyrrolidine dithiocarbamate coprecipitation[5] and analysed by atomic absorption spectrometry using a heated graphite atomiser (E. A. Boyle and J. M. Edmond, unpublished). The estimated precision and accuracy of the analysis is ± 0.1 nmol kg^{-1}. The cadmium data and relevant hydrographic information are listed in Table 1.

Depth profiles of cadmium are shown in Fig. 1 along with profiles of the nutrients phosphate and silicate. It is readily apparent that cadmium is depleted in the surface ocean relative to the deeper waters. Such a distribution indicates uptake by organisms at the surface and regeneration from sinking biological debris deeper in the water column. The range in concentration is over a factor of at least ten, from 0.1 ± 0.1 to 1.1 ± 0.1 nmol kg^{-1}.

The cadmium profiles resemble those of phosphate more than silicate: the concentrations increase rapidly with depth to a maximum at ∼ 1 km and then decrease slightly below this. The covariance with phosphate is consistent for all stations (Fig. 2) and suggests that cadmium is regenerated in a shallow cycle, like the labile nutrients, rather than deeper in the ocean, as is silicate.

The results are compatible with some recent cadmium data. Bender and Gagner[4] report values for the Sargasso Sea; surface (0.05 nmol kg^{-1}) and deep (0.30 nmol kg^{-1}): Knauer and Martin[6] give a value for the 1,000-m depth near Hawaii (1 nmol kg^{-1}) and report surface concentrations below their detection limit (0.2 nmol kg^{-1}). Estimates of the phosphate concentrations at these locations were made using data from nearby GEOSECS stations, and are plotted in Fig. 2. The agreement is excellent. The cadmium–phosphate correlation observed in these stations probably holds throughout the world ocean.

A detailed comparison of the chemical composition of plankton with the hydrographic correlation of cadmium and phosphate is not possible because few analyses include both elements. Goldberg *et al.*[7] report analyses of *Sargassum* and planktonic animals with Cd/P molar ratios of 4×10^{-4} and 6.5×10^{-4} respectively. These values are close to the water column ratio 3.5×10^{-4}. Martin and Broenkow[8] observed

Table 1 Cadmium and associated hydrographic data from three Pacific GEOSECS Stations

Depth (m)	Potential temperature (°C)	Salinity (‰)	Silicate (nmol kg^{-1})	Phosphate (nmol kg^{-1})	Cadmium (nmol kg^{-1})
Station 219 10/8/73	177°17.5'W, 53°6.6'N				
5	7.107	33.069	32.2	1.60	0.61
160	3.533	33.524	74.2	2.42	0.96
349	3.548	33.886	100.8	2.89	1.13
460	3.417	34.027	113.7	2.97	0.96
599	3.294	34.159	128.0	3.03	1.14
642	3.226	34.177	131.0	2.89	1.04
792	3.047	34.266	141.9	2.89	1.05
943	2.834	34.337	150.9	3.04	1.16
1,189	2.531	34.413	163.3	2.95	1.09
1,440	2.249	34.480	176.5	3.05	1.05
1,677	1.983	34.536	190.1	3.01	1.15
1,988	1.741	34.585	202.7	2.98	1.04
2,287	1.594	34.613	209.5	2.91	1.08
2,583	1.482	34.635	214.6	2.89	1.08
3,480	1.310	34.665	226.9	2.80	1.03
3,711	1.281	34.669	221.7	2.74	0.99
Station 226 11/9/73	170°36.5'E, 30°34.0'N				
8	24.73	35.055	3.3	0.02	0.12
36	24.75	35.053	2.7	0.01	0.09
82	23.58	35.005	3.2	0.02	0.20
207	15.19	34.627	7.2	0.49	0.13
382	12.01	34.396	16.0	0.89	0.27
456	10.33	34.277	23.8	1.15	0.20
555	7.92	34.103	38.5	1.61	0.36
591	7.28	34.066	45.6	1.76	0.54
687	5.62	34.031	68.5	2.29	0.89
982	3.62	34.265	121.6	2.97	0.89
1,126	3.20	34.371	135.3	3.02	1.08, 1.06
1,274	2.85	34.440	145.0	3.03	0.96, 0.89
1,421	2.53	34.488	153.7	3.00	0.92
1,712	2.06	34.560	165.1	2.89	0.88
2,442	1.48	34.639	163.5	2.67	0.61
3,005	1.28	34.663	154.7	2.58	0.61
3,785	1.141	34.679	149.7	2.47	0.65
4,175	1.095	34.685	147.9	2.43	0.78
4,361	1.075	34.684	145.7	2.43	0.81
4,779	1.03	34.689	139.9	2.36	0.60
5,446	0.971	34.695	136.2	2.33	0.78
Station 293 3/1/74	178°5.0'W, 52°40.0'S				
3	11.328	34.419	2.7	0.72	0.13
216	7.65	34.419	5.2	1.28	0.31, 0.38
458	6.46	34.342	6.6	1.40	0.20
685	5.27	34.268	17.6	1.76	0.46
961	4.07	34.312	30.8	2.00	0.71, 0.73
1,143	3.30	34.347	44.8	2.13	0.98
1,290	2.83	34.395	53.8	2.24	0.91
1,587	2.45	34.525	68.9	2.28	0.83
1,733	2.37	34.577	73.4	2.27	0.87
2,029	2.19	34.659	80.3	2.19	0.78
2,326	2.018	34.708	83.7	2.10	1.07, 1.12, 1.29
2,621	1.78	34.732	88.5	2.06	0.96
2,771	1.65	34.763	91.9	2.06	0.61
3,065	1.422	34.735	98.3	2.08	0.66
3,509	1.05	34.723	108.7	2.10	0.77
4,019	0.77	34.711	116.6	2.13	0.80
4,474	0.586	34.705	120.5	2.15	0.82
4,926	0.48	34.700	124.5	2.16	0.98, 1.26
5,271	0.446	34.699	125.3	2.17	0.92

cadmium enrichments by a factor of 3 in plankton from the Baja California upwelling region relative to organisms in the open Pacific. Cadmium, like phosphate, is enriched in the surface waters of upwelling regions (for example, Stn 219) relative to areas of lower productivity. The phosphate content of the Baja California waters is about a factor of 3 higher than further away from the coast (R. C. Dugdale and M. L. Healy, unpublished; F. R. Richards, unpublished), and a similar enrichment factor should apply to the cadmium concentration in these waters. It is evident that the cadmium content of plankton depends on the concentration in the water in which the organisms grow. A mechanistic interpretation of the cadmium–phosphate correlation requires information on the site of incorporation of cadmium in organisms and sinking biogenic debris. Similar depth profiles would result if cadmium were carried in either the soft tissues of organisms or in the more

rapidly dissolving hard parts, such as celestite ($SrSO_4$) (ref. 9). Future work on the chemical composition of marine plankton should attempt to establish the location of the various elements in the organisms.

There is little information on the cadmium content of river waters. Where measurements have been attempted, concentrations were often near or below the detection limits of the instrumentation (ref. 10 and refs therein). To derive an estimate for the continental supply rate of this element, a glass fibre-filtered sample from the mouth of the Amazon River was obtained at Macapa, Brazil in May 1974. The cadmium concentration of this sample (0.6 nmol kg^{-1}) may be tentatively taken as representative of world rivers. The resulting estimate of the oceanic 'residence time' of cadmium is ~ 50,000 yr.

Last, it should be noted that there is no evidence in these data for cadmium pollution in the open ocean. Indeed, the

efficient scavenging of this metal by organisms will rapidly remove pollutant cadmium from the mixed layer so that it should not be expected to accumulate in the surface ocean.

E. A. Boyle
F. Sclater
J. M. Edmond

Department of Earth and Planetary Sciences,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

1 Brewer, P. G., in *Chemical Oceanography*, 2nd ed., 1 (edit. by Riley, J. P., and Skirrow, G.), ch. 7, 415 (Academic, London and New York, 1975).
2 Boyle, E. A., and Edmond, J. M., *Nature*, 253, 107–109 (1974).
3 Sclater, F., Boyle, E., and Edmond, J. M., *Earth planet. Sci. Lett.* (in the press).
4 Bender, M. L., and Gagner, C. L., *J. mar. Res.*, 31, 119–128 (1976).
5 Boyle, E. A., and Edmond, J. M., in *Analytical Methods in Oceanography*, ACS Advances in Chemistry Series, No. 147, 44–45 (ACS, Washington, DC, 1975).
6 Knauer, G. A., and Martin, J. H., *Limnol. Oceanogr.*, 18, 597–604 (1973).
7 Goldberg, E. D., Broecker, W. S., Gross, M. G., and Turekian, K. K., in *Radioactivity in the Marine Environment*, 144 (National Academy of Sciences, 1971).
8 Martin, J. M., and Broenkow, W. W., *Science*, 190, 884–885 (1975).
9 Brass, G. W., and Turekian, K. K., *Earth planet. Sci. Lett.*, 23, 141–148 (1974).
10 Wakita, H., in *Handbook of Geochemistry*, II-2 (edit. by Wedepohl), 48–I–1 (Springer, Berlin, 1970).

Combustion sources of atmospheric chlorine

Estimates of the primary sources of tropospheric gaseous chlorine compounds (ClX) which have the potential for reaching the stratosphere, list man-made halogenated hydrocarbons (such as refrigerants and aerosol propellants), man-made and natural hydrogen chloride sources (such as burning of waste materials and volcanic activity), sea spray and naturally occurring chlorinated hydrocarbons. The existence of an unidentified source of tropospheric ClX has been proposed[1-4] which has since been postulated to be combustion. The purpose of this paper is to examine and estimate the strength of fires as ClX sources.

Chlorine, as well as the other halogens, serve as free-radical traps in flames and are effective in suppressing flaming combustion and promoting glowing combustion[5]. When chlorine compounds are heated in contact with cellulose materials, ClX is produced either by reactions involving the hydroxyl groups of the cellulose, or by reactions involving 'active' water formed *in situ* by dehydration[6]. About 1% of the combustion products evolved from wood is methane[7], which can be converted to methyl chloride by the chlorine available in the wood. Bethge and Troeng[8] found that the natural chlorine content of wood pulp and other cellulosic materials ranged from 0.07 to 2.68 mg g^{-1}, that the chlorine content of wood was greater by an unspecified amount than that of unbleached wood pulp, and that 86% of the chlorine content of the sample was lost during combustion.

Lovelock (personal communication) refined his earlier estimate[1] to 1 cm^3 of methyl chloride gas produced for each gram of cellulose burned in glowing combustion. This is approximately equivalent to 2.2 mg of methyl chloride or 1.6 mg Cl produced per gram of cellulose burned. If the 86% loss of Cl during burning is representative, the average initial chlorine level in cellulose is 1.8 mg g^{-1}. There is good agreement between experiments, excepting those which involved flaming combustion.

The common combustion products of most burning plastics are smoke, CO, CO_2 and H_2O vapour. Trace amounts of extremely complex liquid, solid and gaseous combustion products also occur[9]. PVC is different from other plastics because its principal combustion product is HCl (refs 10, 11). The overall chlorine content of PVC is ~ 49%, varying with the amount of various pigments and lubricating organic additives. Yields of HCl during thermal decomposition vary from almost complete (583 of a possible 584 mg (ref. 12)) to ~ 95% (ref. 10).

PVC is widely used for many purposes and consequently is usually involved to some extent in any man-made fire. The most reliable estimates of PVC production in the USA are given in Table 1. In 1974, 6.5% of this production went into various forms of packaging materials[13], which are ultimately disposed of as waste. As a result of Executive Order 507, much of the municipal waste and waste from suburban and rural areas in the USA is now buried. Some of the PVC is burnt in free-burning fires such as those in fireplaces and home incinerators. Much of the HCl produced in these fires does not reach the atmosphere because it is quickly deposited on adjacent surfaces, is mixed with smoke particles (both solid and liquid) and is absorbed by the water produced during combustion. An indication of the reactivity of HCl and its presence in smoke from the burning of waste is the heavy corrosion of the metal parts of municipal incinerators and the observed inhibition of combustion of municipal wastes[14]. I estimate that ~ 50% of the packaging material was burnt in 1965, decreasing linearly to 10% in 1974, and that 10% of the evolved Cl is converted to methyl chloride and other chlorinated hydrocarbons. Table 1 shows that the production of methyl chloride from waste has only recently declined.

Disposition of the rest of the PVC in the USA apparently is not available. English estimates[15] are that 50% of the PVC produced goes into construction. The National Fire Protection Association estimate of the loss rate of new construction because of fire is $0.59 per $100 or ~ 0.6% yr^{-1} (ref. 16). Thus, ~ 0.3% of the PVC produced yr^{-1} is burnt in urban and industrial fires and ~ 10% of this amount is converted to chlorinated hydrocarbons. As shown in Table 1, the total man-made production of methyl chloride is continuing to rise, in spite of the reduction in the burning of waste.

The area burnt in wildfires is well documented[17]. In 1974, 537,020 ha were burnt in the small wildfires. The amount of fuel per unit area (fuel loadings) was quite variable ranging from 1 to 2,500 t ha^{-1}. Estimates derived from wildfire data (O. L. Holmes, US Forest Service, R-5, personal communication) and other sources[16-21] indicate that the average fuel loading is ~ 30 t ha^{-1}. Of this ,~ 40% is fast burning fine material; the residual is coarse material which tends to burn by glowing combustion.

Controlled burning encompasses a large range of activities including forest residue reduction, agricultural waste burning and other mild fires. The weather and fuel conditions for these fires are deliberately chosen to promote low flames and glowing combustion to minimise fire intensity, and to reduce the probability of fire escape. Much of the agricultural residual materials have a high mineral content (for example, rice straw 15%, sugar cane 30%) which also promotes glowing combustion and increases the production of methyl chloride.

Yamate's[22] survey of the burning of forest and agricultural residues in the United States is in agreement with other sources[19-21] except for forest residue reduction in the Pacific North-western States which is 424 times greater. The Forest Service was taken as the authoritative source. All values in Table 2 of the amount of fuel burnt in forests have been decreased by 40% to account for the rapid burning of fine fuels.

It is generally accepted by observers of large fires that the convection columns reach altitudes of 10 km or more[23]. Cramer of the US Forest Service (personal communication) measured the tops of the 1933 and 1939 Tillamook, Oregon, fires at 12 and 13 km. The convection column above the 1967 Sundance Fire in Montana reached 10.5 km and was carried to 13 and 14 km by the isentropic flow field downwind[21].

Table 2 Methyl chloride from combustion sources in the USA, 1972–74

Source	Area burnt (ha)	Fuel loading (t ha^{-1})	Fuel burnt (t yr^{-1})	Methyl chloride (t yr^{-1})	Totals (t yr^{-1})
Agricultural					
Field residues	1,767,771	1.13	2.00×10^6	2,700	2,700
Forests					
Forest residues					
Pacific North-west	Not currently available		39.4×10^6	52,000	
California	Not currently available		0.9×10^6	1,200	
Southern States	8.10×10^5	25	20.0×10^6	26,700	
Rest of USA	8.50×10^5	3	2.5×10^6	3,300	83,200
Wildfires (A, B, C, D, E)	4.69×10^5	30	14.1×10^6	18,600	
Wildfires (F, G)	2.29×10^5	30	18.9×10^6	24,900	43,500
Man-made (Table 1)					80,600
Totals					210,000

These large fires may spread at rates up to several thousand ha h^{-1}. Fuel loadings vary from one to as much as 1,250 t ha^{-1}. Combining the fire spread rates with the heat of combustion of wood (~ 5,000 calorie g^{-1}) gives an energy release rate of 10^9–10^{11} calorie s^{-1}. Such energy release rates are as great as the largest natural thunderstorms[24]. The primary energy source is locally intensive at the ground rather than in the extensive of the cloud as in a thunderstorm. Studies of convection columns[22] show that most of the air passes through or near the bottom of the column in or near the combustion zone. There is a little evidence of lateral entrainment into these convection columns at heights > 100 m above the fire. Condensation within the convection column is apparently determined by the lifting condensation level of the ambient air at the ground level. This type of fire usually occurs in extremely dry conditions and, if condensation occurs at all, the level is > 6,000 m.

Analysis of the Project Flambeau convection column[23] showed that these large fires followed a consistent pattern, and that a fire had to have at least 27 ± 2 hectares burning (not burnt) to be classified as a mass fire. The application of the Project Flambeau results to large wildfires requires an analysis of the geometry of these fires. In large, vigorously spreading wildfires the flame front at the fire's head is ~ 150 m thick in timber and 50 m thick in brush. The shape of the burnt area is a complex function of terrain, wind, humidity, fuel moisture, fuel distribution, percentages of live and dead materials and other factors. A variety of simple models can be assumed, but the simplest is a square, of side 2r, with two semicircles of radius r at each end. The active fire front is of radius r and of thickness d which approximates the final form of the Sundance fire[17].

The total area burnt A_B is the sum of the end semicircles and the centre square

$$A_B = \pi r^2 + 4r^2 \qquad (1)$$

The significant burning area, A_s, is

$$A_s = \pi (2rd + d^2)/2 \qquad (2)$$

Solving equation (2) for r and substituting into equation (1) gives the area burnt during a mass fire as

$$A_B = (\pi + 4)[(A_s/d) + (d/2)]^2 \qquad (3)$$

Substituting the characteristic values for d gives minimum sizes of $A_B = 4.09 \times 10^6$ m^2 for timber and $A_B = 4.16 \times 10^7$ m^2 for brush.

Fires this size are categorised as class F (1,000 to 4,999 acres) and class G ($\geqslant$ 5,000 acres) in the USA where there were 227 class F and 53 class G wildfires in 1974 (ref. 17). These fires burnt a total of 628,605 ha (1,552,654 acres), which is a normal yearly area. The flame front primarily burns the fine fuels which average 12 t ha^{-1}. About 50% of the remaining fuel, or 8 t ha^{-1}, is later burnt in glowing combustion and produces methyl chloride. This air is entrained into the convection column and transported to at least tropopause levels. The results of the various production processes are shown in Table 2.

Taking account only of documented fires, I estimate the production of methyl chloride in the USA by burning cellulose and PVC averages 210,000 t yr^{-1}. The requirement that ClX be produced during glowing combustion may be overly stringent and the estimate may be 40% too low. The rate of production by this source has been estimated to be 5.2×10^6 t yr^{-1} (ref. 26). For instance, current estimates are that half of the timber cut in the world is used for firewood[27]. Forest fire wood and agricultural burning has been a common practice throughout the world for many thousands of years, swamping the US ClX production. If Eckholm's estimates[27] of forest depletion are correct, by the year 2000 there will be a significant decrease

Table 1 Recent US yearly production of polyvinyl chloride and estimated release of methyl chloride

Year	PVC production (Mt yr^{-1})*	Methyl chloride produced from PVC burnt in waste (t yr^{-1})	Methyl chloride from building fires (t yr^{-1})	Total methyl chloride (t yr^{-1})
1965	8.006	26,020	24,018	50,038
1966	9.462	28,018	28,386	56,404
1967	9.435	25,212	28,305	53,517
1968	10.04	23,929	30,120	54,049
1969	13.03	27,291	39,090	66,381
1970	13.29	23,996	39,870	63,866
1971	15.58†	23,630	46,740	70,370
1972	19.60‡	24,064	58,800	82,864
1973	20.08§	18,853	60,240	79,093
1974	21.87‖	14,216	65,610	79,826

in CIX production because of lack of wood and cellulose for burning.

THOMAS Y. PALMER

823 South Ridge Drive
Fallbrook, California 92028

1 Lovelock, J. E., *Nature*, **256**, 193–194 (1975).
2 Lovelock, J. E., *Nature*, **252**, 292 (1974).
3 Lovelock, J. E., *Nature*, **258**, 776 (1975).
4 Cicerone, R. J., Stedman, D. H., and Stolarski, R. S., *Geophys. Res. Lett.*, **2**, 219–222 (1975).
5 Lyon, J. W., *The Chemistry and Uses of Fire Retardants*, 462 (Wiley-Interscience, New York, 1970).
6 Halpern, Y., and Patai, S., *Israel J. Chem.*, **7**, 685–690 (1969).
7 Rasbash, D. J., *Trans. J. Plastics Inst.*, **4**, 55 (1967).
8 Bethge, P. O., and Troeng, T., *Svensk PappTidn.*, **62**, 598–601 (1959).
9 Lipska, A. E., and Alvares, N. J., *Paper 71-2 Western States Section, The Combustion Institute* (1971).
10 Woolley, W. D., *Br. Polym. J.*, **3**, 186 (1971).
11 Gerstle, R. W., and Devitt, T. W., *64th Annual Meeting APCA, Atlantic City, New Jersey* (1971), quoted in ref. 10.
12 Boettner, E. A., Ball, G., and Weiss, B., *J. appl. Polym. Chem.*, **13**, 377–391 (1969).
13 Anon., *Modern Packaging*, 24 (July 1974).
14 Connolly, H. H., *Polym. News*, 1(11/12), 25 (1975).
15 Anon., CP 91/74, *BRE, Fire Res. Station, Borehamwood, England*, 17 (1975).
16 Bugbee, P., *Men Against Fire* (National Fire Protection Association, Boston, Massachusetts, 1972).
17 Anon., 1974 *Wildfire Statistics* (USDA, Forest Service, Washington, DC, 1975).
18 Cooper, R. W., and Mobley, H. E., *Presented at 3rd National Conference on Fire and Forest Meteorology, Lake Tahoe, California*, April 2–4, 1974 (1974).
19 Cramer, O., (ed.), *GTR PNW-24*, PNW Station, USDA, Forest Service, Portland, Oregon (1974).
20 Fahnestock, G. R., *Problem Analysis*, PNW Station, USDA, Forest Service, Portland, Oregon (1974).
21 Finklin, A. I., *GTR INT-6*, INT Station, USDA, Forest Service, Ogden, Utah (1973).
22 Yamate, G., Report EPA–450/3–74–062, NITS, Arlington, Virginia, 25p (1974).
23 Countryman, C. M., *PSW-19*, PSW Station, USDA, Forest Service, Berkeley, California (1964).
24 Sellers, W. D., *Physical Climatology* (University of Chicago Press, Chicago, 1965).
25 Palmer, T. Y., and Northcutt, L. I., *Fire Technol.*, **11**, 111–118 (1975).
26 Ryan, J. A., and Mukherjee, N. R., *Rev. Geophys. Space Phys.*, **13**, 650–658 (1975).
27 Eckholm, E. P., *Losing Ground, Environmental Stress and World Food Prospects* (W. W. Norton, New York, 1976).

Modelling climatic response to orbital parameter variations

ACCORDING to the astronomical or Milankovitch theory, the glacial fluctuations of the Pleistocene are the climatic response to secular variations in the obliquity, eccentricity, and longitude of perihelion of the Earth's orbit. Variations in these orbital parameters during the past several hundred thousand years can be computed with great precision[1], and the resulting changes in incident radiation at the top of the atmosphere are easily obtained. If one can predict the climatic response to perturbations in incident radiation, one can, therefore, test the Milankovitch theory by comparing the predicted climatic changes with those inferred from the geological record. No other theory of the ice ages admits such a straightforward check on its validity. We have made a preliminary attempt at verifying the Milankovitch theory using a zonally symmetric energy-balance climate model forced with seasonally varying insolation and obtain generally favourable results.

The model predicts, as functions of latitude and time of year, the temperatures of two atmospheric layers (representing the upper and lower halves of the troposphere), surface temperatures over land and ocean, the depth of snow over land, and the sea ice thickness. A radiative transfer model similar to that of Manabe and Wetherald[2] is used to compute radiative heating rates in the two atmospheric layers and the net flux at the surface. Cloudiness and relative humidity are assumed to be fixed and independent of latitude in the radiative calculation. Heat is transferred meridionally by linear diffusion in the atmosphere.

This crude atmospheric model is coupled to the land and ocean surfaces through simple boundary layers. The ocean temperature is assumed to be that of a 40-m deep isothermal layer of water, and the land temperature that of a water-saturated surface with no heat capacity. No heat is transported by the model's 'ocean'. Ocean temperatures are not allowed to fall below 273 K, at which temperature sea ice is assumed to form. Sea ice thickness and temperature are predicted with a
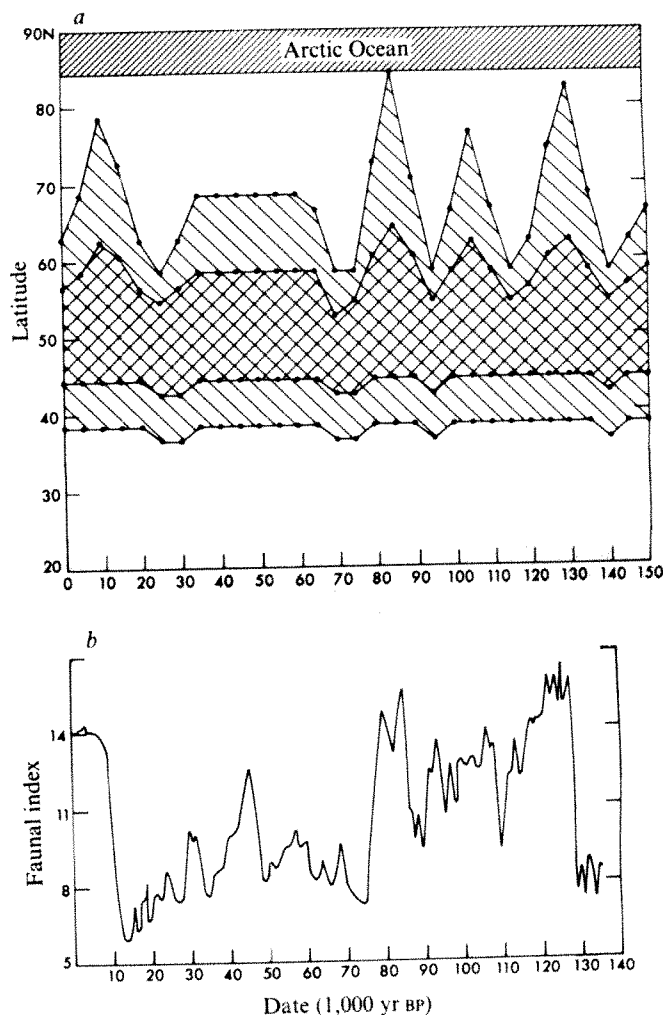
Fig. 1 The model's 'palaeoclimatic record' for the past 150,000 yr is shown in the upper figure. *a*, Four points are plotted for each experiment, corresponding to the seasonal limits of land snowcover (hatching) and sea ice (cross hatching) in the Northern Hemisphere. The northermost point represents the minimum (summer) extent of snowcover on land. North of this latitude, snow is free to accumulate from year to year; we therefore identify it with the extent of continental glaciers. The other three points (from north to south) represent the minimum extent of sea ice, the maximum extent of sea ice, and the maximum extent of snowcover over land. Changes in snowcover and sea ice extent predicted by the model in the Southern Hemisphere are much smaller than these changes in the north. ('Arctic Ocean' refers to latitudes at which there is no land.) *b*, A plot of palaeo-temperatures inferred from faunal abundances (of North Atlantic plankton) in a deep-sea core in the North Atlantic[6]. Its gross features are consistent with other palaeoclimatic reconstructions of this period. Three warm peaks separated by ∼ 20,000 yr mark the last interglacial. The rapid descent into the Wisconsin ice age occurs at ∼ 70,000 yr BP, followed by a period of relatively warm climate (middle Wisconsin), then by the late Wisconsin glacial maximum, and finally by the warmth of the Holocene.

scheme very similar to that used by Bryan[3]. It is also assumed that snow falls at a fixed rate, dependent only on latitude, whenever the surface temperature falls below 273 K. Snowmelt and snowdepth are then computed in a straightforward manner. Surface albedos are chosen to be 0.1 over bare land and open water, and 0.7 over snow-covered land and sea ice. Zonally symmetric atmospheric heating rates are obtained by combining the predicted rates over land and ocean in proportions determined by the fraction of land around a latitude circle.

The model is similar to diffusive energy-balance models used by Sellers[4] and North[5], with changes in surface albedos having a dominant role in determining the sensitivity to changes in external forcing.

For a given set of orbital parameters, we compute the flux of solar radiation incident at the top of the atmosphere as a function of latitude and time of year. The model's equilibrium response to this pattern of insolation is then obtained by integrating its equations of motion until a steady state is achieved. The results of such computations for the present orbital parameters and for those at 5,000-yr intervals into the past (for the past 150,000 yr) are summarised in Fig. 1a. Figure 1b (reproduced from Sancetta, Imbrie, and Kipp[6]) is a plot of a faunal index obtained from counts of planktonic foraminifera in a single deep-sea core in the North Atlantic. The index is calibrated to be an estimate of summer sea-surface temperatures in °C.

The largest difference in mean Northern Hemisphere surface temperature between any two of the experiments described in Fig. 1a is 2.4 K. Changes in mean Southern Hemisphere surface temperatures are smaller—at most 0.4 K.

The model is particularly sensitive to insolation anomalies during the Northern Hemisphere summer, with high obliquity (large seasonal variation) and perihelion near Northern Hemisphere summer solstice producing the warmest climates. During the summer, when land surface temperature gradients are small and incident solar radiation is large, changes in temperature produce the largest displacements of the snow margin, and the resulting changes in albedo produce the largest changes in absorption of incident radiation. The Southern Hemisphere, with less land and a smaller seasonal variation, is not as sensitive to the seasonal redistribution of insolation. This sensitivity to summer insolation in the Northern Hemisphere is consistent with Milankovitch's assumptions, and the time evolution predicted by the model is similar to his summer insolation curves for high latitudes in the Northern Hemisphere[7]. The reader is referred to the literature[8,9] for a detailed comparison of the Milankovitch curves with the geological record.

Large discrepancies between the model and the record of the recent past are evident from Fig. 1. For the present values of the orbital parameters, the model gives a climate intermediate between a predicted glacial minimum at 10,000 yr BP and a glacial maximum at 25,000 yr BP, while observations place the present climate much closer to the most recent climatic optimum at ~ 5,000 yr BP. This discrepancy may arise from other sources of variability, or from a lag resulting from the time required to change the temperatures of abyssal waters or to build and destroy continental glaciers, or to other model deficiencies which may tend to overemphasise the importance of summer insolation anomalies in the Northern Hemisphere.

In spite of such discrepancies, Fig. 1 suggests that a substantial portion of climatic variability on these time scales can be understood as the equilibrium response to perturbations in the orbital parameters. Quantitative tests with more convincing climate models are needed to confirm these very preliminary results. If this theory is correct, the palaeoclimatic record becomes an invaluable source of information on climatic sensitivity.

MAX J. SUAREZ*
ISAAC M. HELD†

Geophysical Fluid Dynamics Program,
Princeton University,
Princeton, New Jersey 08540

* Present address: Department of Atmospheric Sciences, University of California at Los Angeles, Los Angeles, California.
† Present address: Center for Earth and Planetary Physics, Harvard University, Cambridge, Massachusetts.

1 Berger, A. L., *Theorie Astronomique des Paleoclimats* (Université de Louvain, Louvain, 1973).
2 Manabe, S., and Wetherald, R. T. *J. atmos. Sci.*, 24, 241–259 (1967).
3 Bryan, K., *Mon. Weath. Rev.*, 97, 163–172 (1969).
4 Sellers, W. D., *J. appl. Meteorol.*, 12, 241–254 (1973).
5 North, G. R., *J. atmos. Sci.*, 32, 1301–1307 (1975).
6 Sancetta, C., Imbrie, J., and Kipp, N. G., *Quat. Res.*, 3, 110–116 (1973).
7 Milankovitch, M., *Canon of Insolation and the Ice-Age Problem* (Royal Serbian Academy, Belgrade, 1941; translated by Israel Program for Scientific Translations, Jerusalem, 1969).
8 Chappell, J., *Nature*, 252, 199–202 (1974).
9 Broecker, W. S., *et al.*, *Science*, 159, 297–300 (1968).

Ultraviolet absorption by metal–ammonia solutions

By pulse radiolysis of aqueous solutions at room temperature, new absorption bands for H, OH, D, OD and e_{aq}^- in the ultraviolet region down to 188 nm have been observed[1,2] and recently confirmed[3]. The favoured interpretation was a redshift of the absorption band of water caused by perturbation by the solutes (H, D, e_{aq}^-) similar to the β-band displacement in alkali halide crystals[1]. It was also suggested that such a band may exist in liquid ammonia, implying that it is a general phenomenon associated with the solvated electron. Furthermore, the transient spectrum obtained by pulse radiolysis of pure liquid ammonia (-45 °C) shows, in addition to the broad infrared absorption of the solvated electron, two unidentified bands in the ultraviolet region[4]. To study optical properties in this range, we chose metal–ammonia solutions which are relatively stable although their decomposition leads to the amide ion which itself absorbs strongly in the same region. Our findings strongly support the hypothesis, though amide and hydrogen are the only products and have been thoroughly studied.

The solutions were prepared taking care to ensure that the reagents were sufficiently pure to avoid their catalytic decomposition. We used pressure-resistant cells with Suprasil windows[5]. To account for the small difference between the optical paths of the sample and reference cells which could cause some solvent absorption in the ultraviolet region, we systematically recorded the base line with pure ammonia in both cells. To record the spectra for a wide range of alkali metal concentrations, various optical path lengths from 10^{-2} to 10 mm were used.

The temperature was varied by a current of cold nitrogen gas and we checked that thermal equilibrium was reached before recording the spectrum, otherwise the shift of the solvent

Fig. 1 Experimental results for the dependence of the extinction coefficient on wavelength. For potassium–ammonia solutions: ●, -50 °C, 2.35×10^{-3} mol l^{-1}, path length 1.90 mm; ○, 20 °C, 2.03×10^{-3} mol l^{-1}, path length 1.90 mm; ■, -50 °C, 7.3×10^{-4} mol l^{-1}, path length 10.0 mm; □, 20 °C, 6.3×10^{-4} mol l^{-1}, path length 10.0 mm. The results are normalised relative to $\varepsilon_{e^-}{}_{am, 700, 20 °C} = 1,575$ l mol^{-1} cm^{-1} and $\varepsilon_{e^-}{}_{am, 700, -50 °C} = 2,600$ l mol^{-1} cm^{-1}. Dashed curves: amide spectra; *a*, -50 °C; *b*, 20 °C. $\varepsilon_{NH_2^-, 348, 20 °C} = 2,600$ l mol^{-1} cm^{-1}; $\varepsilon_{NH_2^-, 330, -50 °C} = 3,040$ l mol^{-1} cm^{-1}. Results are mean values for the concentration range 5×10^{-4}–2×10^{-3} mol l^{-1}.
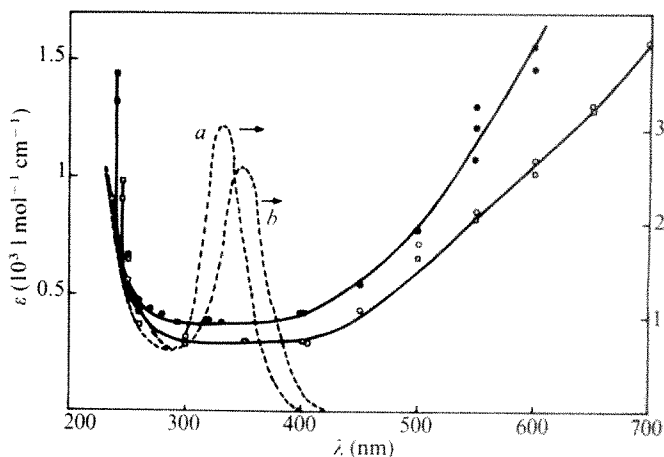
band with temperature results in a pseudo absorption within the studied range. Spectra were recorded with Beckman ACTA IV M and Beckman DK-1A spectrophotometers flushed with nitrogen gas.

At low temperature ($-50\ ^\circ$C) the rate of decomposition is usually negligible but at room temperature the reaction is catalysed by the walls, especially when using the thin cells for which the surface-to-volume ratio is very high. The absorbances were first followed kinetically for a short time in the infrared band of e_{am}^-, for instance at 700 nm, and at the maximum of the amide absorption, $\lambda = 348$ nm (20 $^\circ$C). These measurements revealed the extent to which decomposition occurred during the recording and when necessary allowed the spectra to be corrected.

We examined many more or less stable solutions. The results presented Fig. 1 correspond to the best runs. The optical densities were normalised at 700 nm and the shape of the infrared spectra between 500 and 800 nm at 25 $^\circ$C and between 700 and 1,000 nm at $-50\ ^\circ$C agree with those found previously[4,6].

The spectra normalised relative to A_{700} were plotted assuming an extinction coefficient at 700 nm $\varepsilon_{e^-_{am,700}} = 1,575$ l mol^{-1} cm^{-1} at 20 $^\circ$C, and $\varepsilon_{e^-_{am,700}} = 2,600$ l mol^{-1} cm^{-1} at $-50\ ^\circ$C (refs 4 and 6). It appears (Fig. 1) that the spectrum consists of a second band in the ultraviolet superposed on the tail of the large infrared band and that no characteristic absorption of NH_2^- is present. Only at room temperature and at a potassium concentration of 2×10^{-3} mol l^{-1} was a weak amide band present, but its contribution was easily subtracted since the amide spectrum is accurately known down to 230 nm (Fig. 1)[7,8]. Even when the amide absorption increases with time, the kinetics show that the absorption at 348 nm and 230 nm increases with the decay of the infrared band (isobestic point at 402 ± 2 nm). But the increase in the ultraviolet absorption is not at all that expected from the amide spectrum. This means that a part of the ultraviolet absorption should be attributed to e_{am}^- and so the apparent extinction coefficient ($\varepsilon_{NH_2}^- - \varepsilon_{e_{am}}^-$) enables us to determine corresponding $\varepsilon_{e_{am}}^-$ values in this region. These are in good agreement with those deduced from direct observation of undecomposed solutions (see below). This method of subtracting the amide spectrum is the only one which can be applied to high concentrations ($\sim 10^{-2}$ mol l^{-1}), for which the decay is rapid. In this case the same spectrum was obtained, indicating that ultraviolet absorption occurs when the electron is clearly associated with ion aggregates.

We note, for the best runs presented in Fig. 1, that the results agree for both concentrations, that is 6.3×10^{-4} mol l^{-1} at 20 $^\circ$C (optical path 10 mm) and 2.0×10^{-3} mol l^{-1} at 20 $^\circ$C (optical path 1.9 mm), showing that the ultraviolet absorption is well correlated with the infrared absorption of e_{am}^-. Comparing the spectra of the same solutions at room and low temperature, Fig. 1 shows that the blueshift of the wide infrared band is still observed in the plateau region (300–400 nm) resulting in values

$$\varepsilon_{350,\ 20\ ^\circ C} = 300 \pm 30\ \text{l mol}^{-1}\text{cm}^{-1}$$

and

$$\varepsilon_{350,\ -50\ ^\circ C} = 390 \pm 30\ \text{l mol}^{-1}\text{cm}^{-1}$$

At lower wavelengths, we observe absorptions merging into the solvent band as in the case of e_{aq}^-

$$\varepsilon_{245,\ 20\ ^\circ C} = 950 \pm 100\ \text{l mol}^{-1}\text{cm}^{-1}$$

and

$$\varepsilon_{240,\ -50\ ^\circ C} = 1,400 \pm 100\ \text{l mol}^{-1}\text{cm}^{-1}$$

The ultraviolet absorption from the solvated electron in liquid ammonia increases at $\leqslant 280$ nm while the solvent absorption only starts to rise at $\leqslant 230$ nm. In water the respective values are 220 nm and 180 nm. Although it may not be

decisive, this correlated displacement supports Hart's interpretation that the redshift of the absorption band of the solvent molecules is caused by perturbations attributable to the electron.

This conclusion is strengthened by the observed temperature effect. Indeed, Fig. 1 shows that the ultraviolet absorption of the ammonia solution starts at higher wavelengths at 20 $^\circ$C than at $-50\ ^\circ$C as is the case for the band of the pure solvent itself which exhibits a blueshift of ~ 7 nm when the temperature is lowered over the same range (to be published).

If we admit the above interpretation, the extinction coefficient in the ultraviolet range is not a direct characteristic of the solvated electron but is a parameter determined by the molecules in the solvation shell. In the semi-continuum model[9-11], one must take into account interactions which become weaker with successive layers around the electron. The observed ultraviolet absorption is an averaged sum of the perturbations of transitions of these solvent molecules.

A rough approximation which assures the involvement of a limited number of solvent molecules in the solvation sphere of the electron, four for example, leads to extinction coefficients four times smaller than those indicated experimentally. The band obtained is similar to that of the pure solvent and the shift (~ 27 nm) gives a magnitude of the energy of the perturbation of ~ 0.6 eV.

F. BILLIAU
J. BELLONI

Laboratoire de Physico-Chimie des Rayonnements,
Associé au CNRS, 91405 Orsay, France

E. SAITO

DRA, SRIRMA, CEN Saclay,
91190 Gif-sur-Yvette, France

1 Hart, E. J., Nielsen, S. O., and Pagsberg, P., *J. phys. Chem.*, 73, 3171–3172 (1969).
2 Pagsberg, P., Christensen, H., and Rabani, J., *J. phys. Chem.*, 73, 1029–1038 (1969).
3 Nielsen, S. O., Michael, B. D., and Hart, E. J., (in the press).
4 Belloni, J., Cordier, P., and Delaire, J., *Chem. Phys. Lett.*, 27, 241–244 (1974).
5 Belloni, J., and Saito, E., *Electrons in Fluids* (edit. by Jortner, J., and Kestner, N. R.), 461–471 (Springer-Verlag, New York, 1973).
6 Belloni, J., and Fradin de la Renaudiere, J., *Int. J. Radiat. Phys. Chem.*, 5, 31–39 (1973).
7 Cuthrell R. E., and Lagowski, J. J., *J. phys. Chem.*, 71, 1298–1301 (1967).
8 Corset, J., and Lepoutre, G., *J. chim. Phys.*, 63, 659–662 (1966).
9 Copeland, D. A., Kestner, N. R., and Jortner, J., *J. chem. Phys.*, 53, 1189–1216 (1970).
10 Fueki, K., Feng, D. F., Kevan, L., and Christoffersen, R., *J. phys. Chem.*, 75, 2297–2305 (1971).
11 Fueki, K., Feng, D. F., and Kevan, L., *J. Am. chem. Soc.*, 95, 1398–1407 (1973).

Rainfall characteristics in eastern Sahel

THE recent drought conditions in the Sahel region of Africa have generated studies of long term trends in rainfall and monthly or seasonal atmospheric conditions related to the area[1-5]. The results of these studies, in general, indicate no specific trends in precipitation[3], and some evidence of direct[1] and/or long distance[5] relationships between monthly or seasonal atmospheric state and Sahelian rainfall. We present here the results of a study which suggest the major factor affecting rainfall during the drought was a change in easterly wave activity over northern tropical Africa.

The occurrence and characteristics of easterly waves over northern tropical Africa have been discussed by others[6-8]. In relation to easterly waves and precipitation, Burpee[8] has shown that modulation of rainfall in the region between 0 and 18 $^\circ$W is related to wave passage. He found no clear evidence, however, of a similar wave–rainfall relationship between 0 and 15°E. In the Sudan, Hammer[9] found that spatial and temperal patterns of rainfall approximated the 4–5-d period commonly associated with easterly waves. From evaluation of weather satellite imagery, it has been estimated that 90% of the occurrence of widespread heavy rainfall is directly related to easterly wave passage[10].

To define more clearly the characteristics of Sudanese rainfall, daily cumulative rainfall totals from 200 stations
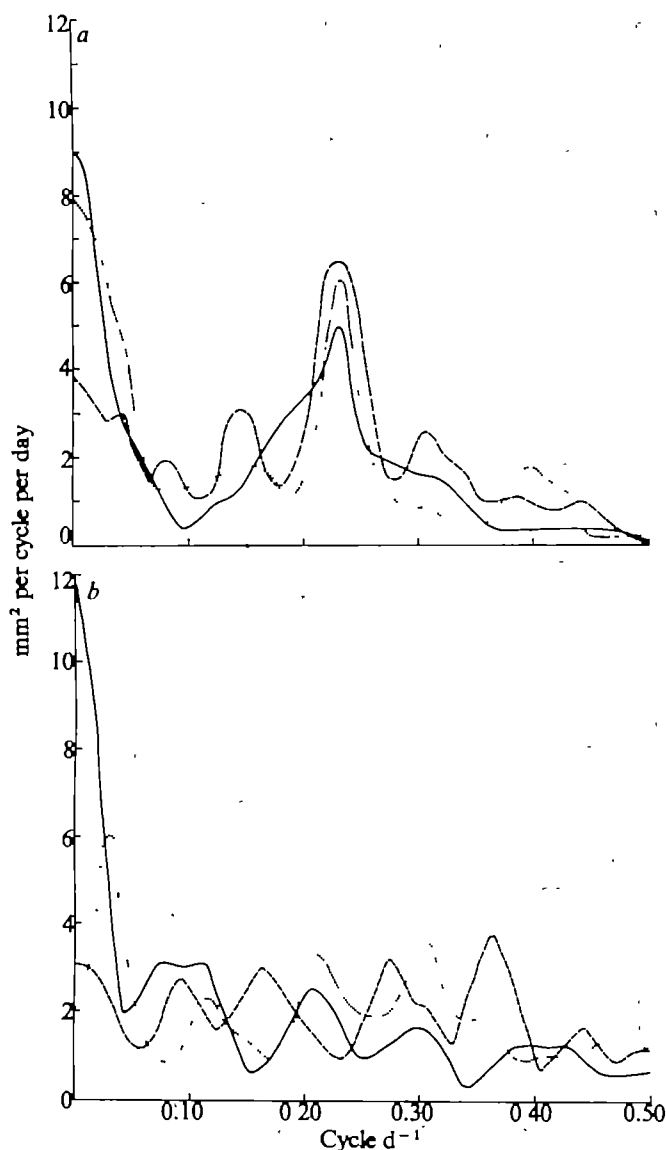
Fig. 1 *a*, Power spectrum analysis results for 90-d period June 13–September 10, 1953 (—), 1961 (·····) and 1969 (————). *b*, As in *a* for 1968 (·····), 1970 (——) and 1971 (————).

in the Sudan were analysed using power spectral techniques. About 75%, a higher proportion in the latter years, were stations located north of 11°N or zonally equivalent to the main drought area of Africa. Unfiltered precipitation totals between June 13 and September 10—the main portion of the rain season—for 1953, 1961, 1968, 1969, 1970 and 1971 were used in the analyses. The spectral analyses were carried out using the OS-3 ARAND system routine SPECT 1 of Oregon State University, and the results are shown in Fig. 1*a* and *b*. The bandwidth used was 0.031 cycle d⁻¹.

Figure 1*a* presents the spectral curves for the years 1953, 1961 and 1969. 1953 and 1961 were years of average to above average rainfall for most of the Sudan during the months of study. The factor accounting for most of the variance in the spectra of 1953 and 1961 was the movement of the so-called Intertropical Convergence (ITC) as shown by the high values at 1 cycle per 90 d (0.011 cycle d⁻¹). For these two years, major varience is also contained within the period 3.7–5 d (0.27–0.20 cycle d⁻¹). This corresponds to the most frequently cited period associated with easterly wave activity. 1969 had a somewhat enigmatic rain season in that in spite of being in the drought sequence, rainfall was near normal in the Sudan. This was similar to precipitation conditions during 1969 in western and central Africa, as found[3,4,11] from integrated station data. The

small amount of variance contained under the 1 cycle per 90-d portion of the curve for 1969 probably reflects the stagnation of the ITC in July, which limited rainfall over the northern half of the Sudan. Because of the restricted effect of the ITC on the 1969 spectrum, caution must be used in interpreting the variance accounted for under the 3.7–5-d portion of the spectrum as being as relatively significant as it seems compared with 1953 and 1961. A frequency domain test was carried out using the OSU OS-3 ARAND system routine NOIZT to determine if each of the time series could be considered 'white noise'. All five peaks on Fig. 1*a* were significant at the 95% level, indicating they were not the result of a purely random process.

In comparison with the three years already discussed, 1968, 1970 and 1971 display a very different spectral pattern (Fig. 1*b*). Although ITC movement in 1970 was statistically significant (evaluated by the NOIZT routine), the remainder of the spectra for the three years are notably less peaked and temporally coincident than those of Fig. 1*a*, with the exception of a possible 20–40-d period in 1968. Rainfall amounts recorded in the Sudan were markedly below average during 1968, 1969 and 1970. Of the nine months, which comprised most of the 90-d period analysed during these three years, five months had 51–75% of the stations with below average monthly rainfall and two additional months with 46–50% of the stations reporting less than average precipitation. The overall appearance of Fig. 1*b* presents a more random peak occurrence pattern than Fig. 1*a*, particularly in the shorter periods. The rainfall of the three dry years thus seems to be responding to a fluctuating rainfall producing mechanism, in contrast to the relatively constant short period pulses of the more humid years.

Visual evaluation of weather satellite imagery of July and August for 1969, 1970 and 1971 confirms the variability of wave conditions over north-eastern tropical Africa (unpublished). Easterly wave-associated cloud pattern passage intervals were notably different during the three years. 1969 had a reasonably consistent 4–5-d period between wave occurrences; in 1970, most waves were spaced at 2–3 d intervals, with an occasional intervening 5-d period between waves: and 1971 was the most erratic year in terms of the period of the occurrence of the wave. Wave periods ranged from 2–9 d, with 2-d spacing most common. For all three years ~ 65% of the waves were conservative for 1–4 d and could be followed from over the Sudan to central Africa, where the satellite tracking used in this study terminated.

Within the Sudan, the drier rain seasons seem to experience rainfall events in a more spasmodic fashion than wetter rain seasons, and major rainfalls are related to easterly wave passage. In addition, the Sudan lies 'upstream', in terms of easterly wave activity, of the western and central African drought areas and there has been generally synchronous rainfall during recent years between the various regions. Thus, there seems to be a strong probability that during the recent drought the easterly wave operated in an unusual manner, with a consequent reduction in rainfall. The intra- and/or extra-wave factors related to modification of wave conditions over continental Africa are still in question.

RICHARD M. HAMMER

Geography Department,
Mount Gravatt College of Advanced Education,
Mount Gravatt, Queensland 4122,
Australia

[1] Bunting, A., Dennett, M., Elston, J., and Milford, J., *Nature*, 253, 622–623 (1975).
[2] Winstanley, D., *Nature*, 243, 464–465 (1973).

3 Bunting, A., Dennett, M., Elston, J., and Milford, J., *Q. Jl R. met. Soc.*, **102**, 59–64 (1976).
4 Tanaka, M., Weare, B., Navato, A., and Newell, R., *Nature*, **255**, 201–203 (1975).
5 Namias, J., *Int. Trop. Meteorology Mtg., Preprint Vol. (Pt 1)*, 141–144 (Am. Met. Soc., Boston, 1974).
6 Carlson, T., *Mon. Weath. Rev.*, **97**, 256–276 (1969).
7 Frank, N., *Mon. Weath. Rev.*, **97**, 130–140 (1969).
8 Burpee, R., *J. atmos. Sci.*, **31**, 1556–1570 (1974).
9 Hammer, R., *J. trop. Geog.*, **35**, 40–50 (1972).
10 Hammer, R., *Geoforum*, **14**, 11–18 (1973).
11 Hastenrath, S., *J. atmos. Sci.*, **33**, 202–215 (1976).

Deep-sea bottom photographs show that benthic organisms remove sediment cover from manganese nodules

MARINE geologists, geochemists, biologists and other oceanographers have studied manganese nodules since they were first reported by Murray and Renard[1] in 1891. Much has been learned (see ref. 2 for a recent review) but many

Fig. 1 Bottom photographs from the manganese nodule province of the North Equatorial Pacific Ocean. A polyplacophoran (*Chiton*) has removed sediment cover from through nodules (*a*) and has almost cleaned the nodule it is on arrowed. Note the sediment covering adjacent nodules. Another *Chiton* (*b*) is feeding on nodule surface. Average nodule in this area is 6–8 cm in maximum length.
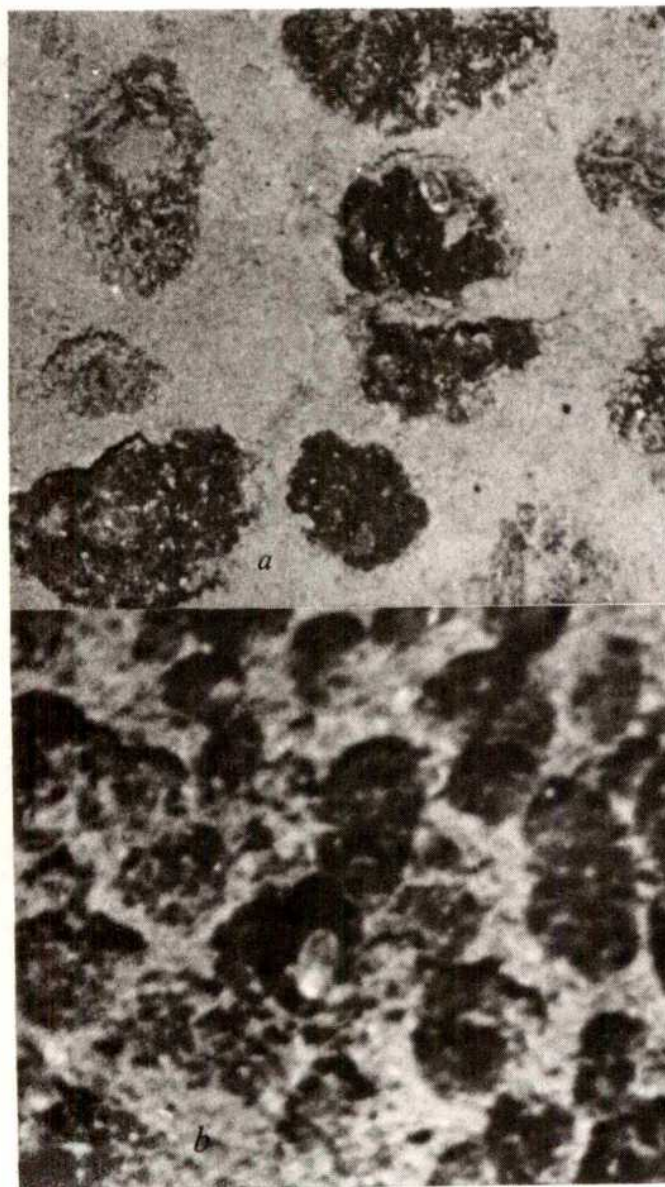
aspects of their formation and growth are open to question. The Environmental Research Laboratory of the National Oceanic and Atmospheric Administration has sponsored a multi-university, multi-disciplinary investigation to determine the environmental impact of manganese nodule mining. As part of the benthic baseline study of the Deep Ocean Mining Environmental Survey (DOMES), approximately 10,000 bottom photographs have been taken. I wish to report a close examination of 365 prints taken on one lowering of the camera near latitude 14°17.5′N, longitude 126°15.4′W at a depth of about 4,500 m.

The following epibenthonts were recognised: 99 cidaroid echinoids; 45 ophiurad ophiuroids; 32 elasipod holothurians; 21 actiniarid zoantharia; 8 pennatulid anthozoans; 5 bryozoans; 4 gorgonid anthozoans; 4 asteroids; 3 porifera; 3 lepidopleurinid polyplacophora; 2 pycnogonida; 2 brachyuran crabs, and one each crinoid, ascidacean and gastropod. The area of each photograph was calculated using the average size of nodules collected at that station as a scale. The average size was 7.4 cm (± 0.7), and it was determined that 3,031 m^2 (± 303) were seen, taking into account the few frames where overlapping occurred. Consequently the epifaunal density is 0.08 organisms per m^2 (± 0.008), which agrees with photographic survey conducted under the oligotrophic Sargasso Sea at similar depths[3].

Animal activities such as burrowing under nodules and pushing them upward or ingesting nodules and subsequently egesting them have been suggested[4] as methods by which the geologically older nodules are kept on top of the younger surficial sediments. These operations may well take place and I would now like to add the mechanism whereby some benthic animals remove sediment from the surface of manganese nodules, apparently as part of their feeding activity. Almost every frame shows evidence of uneven sediment cover on adjacent nodules. In seven frames there are organisms which I believe are grazing and are associated with cleaner nodules (that is, covered with less sediment). The clearest of these photographs are presented here (Figs 1–3). Polyplacophora (chitons) are commonly found in the rocky, intertidal zone where they eat by scraping algae and other material from the rocks[5]. Since manganese nodules are commonly covered with structures of biological origin[6,7] it is not difficult to hypothesise that chitons in a nodule province feed by removing organic material from nodule surfaces (Fig. 1*a* and *b*). Phanerozonid asteroids

Fig. 2 Bottom photograph of phanerozonid asteroid (starfish) feeding on and cleaning manganese nodule surface.
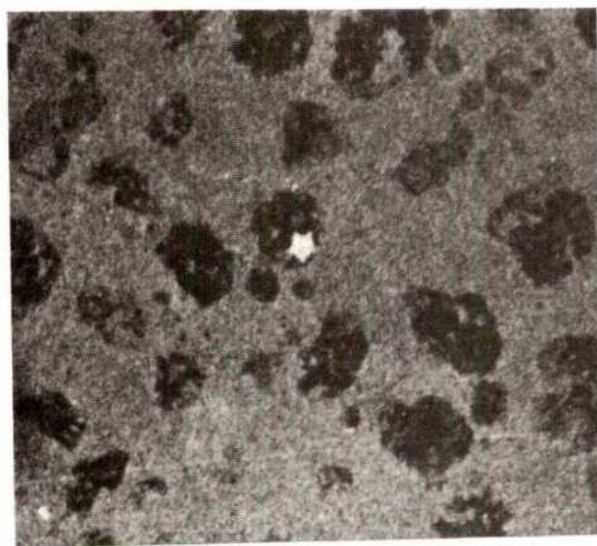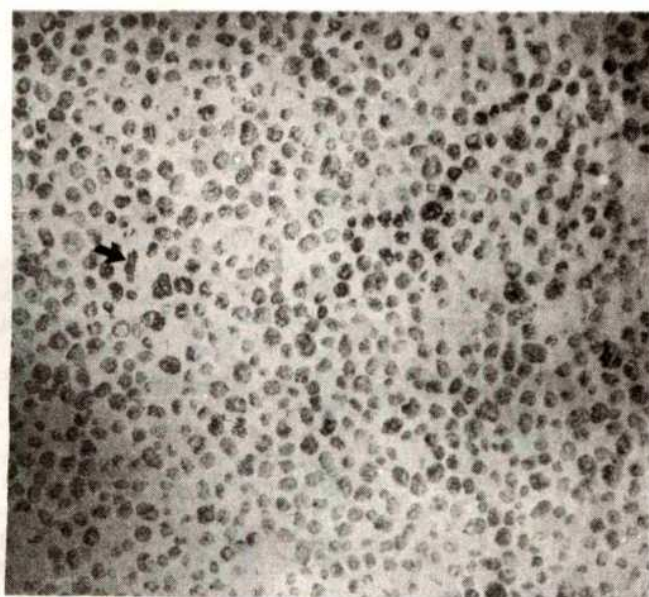
Fig. 3 Bottom photograph of elasipod holothurian (sea cucumber, arrowed). The meandering trail of darker nodules that starts in the upper right corner and ends at the holothurian indicates its path and the removal of sediment cover from nodules as it feeds on deposits.

(Fig. 2) eat various foods[5] and also seem to scrape and clean nodules. Elasipod holothurians are deposit feeders (Fig. 3) and in addition to possibly nudging nodules and keeping them at the sediment–water interface, they ingest sediment from the nodule surface.

Although the sedimentation rate in this area is low[9], sediment redistribution is continually taking place. It is not possible to estimate the rate of either cleaning or bottom redispersion of sediment, but it is evident from these photographs that periodic cleaning takes place. If nodules grow on their upper surfaces this episodic cleaning may help to explain their intermittent growth.

I thank W. S. Broeker and P. Biscaye for discussions, and the Environmental Research Laboratory of NOAA for support.

ALLEN Z. PAUL

Lamont-Doherty Geological Observatory of
　Columbia University,
Palisades, New York 10964

1 Murray, J., and Renard, A. F., *Deep Sea Deposits, Report of the Scientific Results of HMS Challenger, 1873–1876* (1891).
2 Glasby, G. P., *Oceanogr. mar. Biol. A. Rev.*, **12**, 11–40 (1974).
3 Owen, D. M., Sanders, H. L., and Hessler, R. R., in *Deep-Sea Photography* (edit. by Hersey, J. B.), 229–231 (Johns Hopkins Press, Baltimore, 1967).
4 Menard, H. W., *Marine Geology of the Pacific* (McGraw-Hill, New York, 1964).
5 Hyman L. H., *The Invertebrates, Mollusca* (McGraw-Hill, New York, 1967).
6 Greenslate, J., *Nature*, **249**, 181–183 (1974).
7 Wendt, J., *Spec. Publs int. Ass. Sediment.*, **1**, 437–447 (1974).
8 Hyman, L. H., *The Invertebrates, Echinodermata* (McGraw-Hill, New York, 1955).
9 Goldberg, E. D., and Koide, M., *Geochim. cosmochim. Acta*, **26**, 417–450 (1962).

Sulphonium analogue of lecithin in diatoms

PHOSPHATIDYL choline (lecithin) is the most widespread of the phospholipid membrane components of eukaryotic cells, being present in virtually all those investigated so far. The exceptions noted until recently have been blue-green algae[1], the yeast-like fungus *Pullularia pullulans*[2] and the phytoflagellate, *Ochromonas danica*[3]. We report here a sulphonium analogue of phosphatidyl

choline which completely replaces lecithin in the non-photo-synthetic diatom, *Nitzschia alba*[4]. The new phospholipid is a phosphatidyl *S,S*-dimethyl mercaptoethanol (phosphatidyl sulphocholine) with the following structure:

$$
\begin{array}{l}
\overset{\overset{O\;O^-}{\underset{||}{\diagup}}}{H_2C-O-P-OCH_2CH_2S(CH_3)_2^+} \qquad R_1, R_2 = \text{acyl chains}\\[2mm]
\;\;\mid\\
H-C-O-CO-R_1\\[2mm]
\;\;\mid\\
H_2C-O-CO-R_2
\end{array}
$$
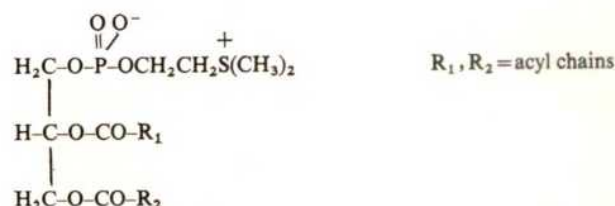
The lipid was first observed on thin-layer chromatograms of the total lipids of *N. alba* (grown in synthetic seawater[4]) as a major component with mobilities and staining behaviour[5] similar to those of lecithin (Fig. 1). But no glycerophosphorylcholine was detected on mild alkaline deacylation of the total lipids, indicating that lecithin was absent in this organism. The new lipid appeared to be a phosphosulpholipid by its labelling with both ^{35}S and ^{32}P when cells were grown in medium containing ^{35}S-sulphate and ^{32}P-phosphate.

The phosphosulpholipid was isolated in pure form by column chromatography of the total lipids on Biosil A silicic acid, using increasing concentrations of methanol in chloroform as eluting solvents. The sulpholipid appeared in the chloromethanol 1:3 eluate and after precipitation from chloroform by addition of acetone, it was chromatographically pure. It accounted for 68% of the total lipid phosphorus and had elemental analyses consistent with those expected for a sulphonium analogue of lecithin. We found C, 64.18%; H, 9.25%; N, 0.00; S, 4.25%; P, 3.85%, atomic ratio P/S, 0.94. The calculated values for $C_{41.9}H_{73.8}O_8SP$ (768.73) are C, 65.47%; H, 9.68%; S, 4.17%; P, 4.02%, P/S, 1.00.

The analysis showed the complete absence of nitrogen and the presence of sulphur and phosphorus in an atomic ratio close to 1:1. Fatty acid analysis by gas–liquid chromatography (GLC) on 10% butanediol succinate polyester at 180 °C revealed the presence of 14:0 (30%), 16:0 (8%), 18:1 (12%), 18:2 (8%), 20:5 (27%) and 22:6 (4%) acids, permitting calculation of the molecular formula given above.

The nuclear magnetic resonance (NMR) spectrum of the

Table 1 Proton NMR assignments for phosphatidyl sulphocholine from *N. alba* in CDCl$_3$

Proton assignment	Phosphatidyl choline* δ (p.p.m.)	Phosphatidyl sulphocholine δ (p.p.m.)	No. of protons calculated	found
C$\underline{H}$=C$\underline{H}$	5.38	5.38	7.5	7.7
$\underline{H}$C–O–	5.1	5.1	1	0.9
C$\underline{H}_2$O–CO–R	} 4.36	4.36	4	4.3
C$\underline{H}_2$O–P (chlorine)				
C$\underline{H}_2$O–P (glycerol)	3.9	3.9	2 }	} 4.3
C$\underline{H}_2$S	—	3.9	2	
C$\underline{H}_2$N	3.8	—	–	—
$\overset{+}{N}$(C$\underline{H}_3$)$_3$	3.37	—	–	—
$\overset{+}{S}$(C$\underline{H}_3$)$_2$	—	3.17	6	5.7
=C–C$\underline{H}_2$–C=	2.83	2.83	5.5	6.0
C$\underline{H}_2$CO	2.28	2.28	4	4.6
C$\underline{H}_2$C=	2.08	2.08	4.1	4.5
C$\underline{H}_2$	1.28	1.28	24	22
C$\underline{H}_3$	0.9	0.9	6	6.0†

*Purified egg lecithin (Serdary Research Laboratories). For assignments see ref. 7.
†Taken as reference for integration.
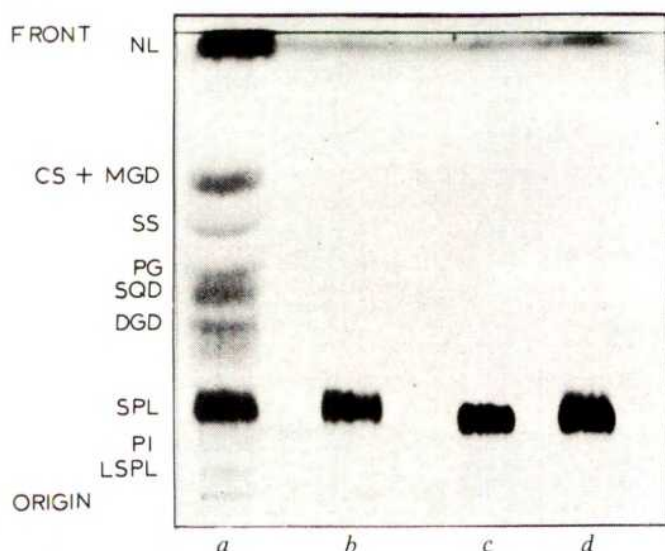
Fig. 1 Thin-layer chromatogram on silica gel H in the solvent system chloroform–methanol–28% ammonia (65:35:5, v/v) of: *a*, total lipids of *N. alba*; *b*, isolated phosphatidyl sulphocholine from *N. alba*; *c*, commercial egg lecithin (PC), and *d*, mixture of phosphatidyl sulphocholine and lecithin. Chromatogram was charred with 50% H_2SO_4. Phospholipids were detected by the phosphate stain[5] and "choline-"containing lipids by the Dragendorff reagent[5]. NL, Neutral lipid; CS, ceramide sulphonate[6]; MGD, monoglycosyl diglyceride; SS, sterol sulphate[6]; PG, phosphatidyl glycerol; SQD, sulphoquinovosyl diglyceride; DGD, diglycosyl diglyceride; SPL, phosphatidyl sulphocholine; PI, phosphatidyl inositol; LSPL, lysophosphatidyl sulphocholine. SPL and LSPL as well as lecithin gave positive tests with the phosphate and "choline" stains.

close similarity except for the sharp peak at 970 cm^{-1} which was absent from the spectrum of the sulpholipid and which is generally assigned to the $(CH_3)_3N^+$ group of lecithin[8,9].

Enzymatic hydrolysis of the sulpholipid with cabbage phospholipase D yielded: (1) a chloroform-soluble phospholipid identified as phosphatidic acid by cochromatography on silica gel H with an authentic standard (R_f, 0.05 and 0.9 in chloroform–methanol–28% ammonia (65:35:5, v/v) and chloroform–methanol–90% acetic acid (30:4:20, v/v), respectively); and (2) a water-soluble sulphur-containing compound which gave a positive Dragendorff reaction[5] and had mobilities on paper chromatography and paper electrophoresis (Table 2) identical to those of synthetic 2-hydroxyethyl dimethylsulphonium iodide (*S,S*-dimethyl mercaptoethanol, "sulphocholine", made by methylation of mercaptoethanol with methyl iodide). The latter product was also present in the 2.5% methanolic-HCl hydrolysate of the sulpholipid along with fatty acid methyl esters and two water-soluble periodate–Schiff-positive products identified chromatographically as α-glycerophosphate and methyl-α-glycerophosphate. Identification of the sulphur-containing acid-hydrolysis product as "sulphocholine" was confirmed by monodemethylation of the sulphonium group with sodium benzenethiolate[10]. The reaction products had GLC retention times relative to dimethylsulphide (4.59 and 6.04) identical to those of authentic hydroxyethylmethyl sulphide and phenylmethyl sulphide, respectively, on a column of 10% butanediolsuccinate polyester on Gaschrom W at 128 °C and a carrier inlet pressure of 0.4 kg cm^{-2}. An identical gas chromatographic tracing was obtained after monodemethylation of synthetic "sulphocholine".

Hydrolysis of the sulpholipid with phospholipase C from *Clostriduim perfringens* released: (1) a chloroform-soluble product with R_f value identical to that of standard 1,2-diglyceride (R_f, 0.22 on silica gel H in petroleum ether–diethyl ether–acetic acid 70:30:1, v/v); and (2) a water-soluble phosphorus and sulphur-containing compound with mobilities similar to those of phosphorylcholine (Table 2). Unlike phosphorylcholine, however, the water-soluble product was found to undergo methanolysis in methanolic 0.2 N NaOH at 37 °C for 1 h, with quantitative release of inorganic phosphate and formation of methoxysulphocholine (2-methoxyethyl dimethylsulphonium salt) identified chromatographically and electrophoretically in comparison with an authentic standard (Table 2). The water-soluble product formed by phospholipase C hydrolysis is thus probably phosphorylsulphocholine.

Mild alkaline deacylation of the phosphatidyl sulphocholine in methanolic NaOH also yielded methoxysulphocholine, together with α-glycerophosphate (Table 2) instead of the expected glycerophosphoryl sulphocholine. Apparently the phosphate-sulphocholine linkage both in phosphatidyl sulpho-

sulpholipid showed a sharp signal at 3.17 δ p.p.m. which was assigned to the six protons of the dimethyl sulphonium group. For comparison, the quaternary ammonium group methyl protons of phosphatidyl choline gave a singlet at 3.37 δ p.p.m. (Table 1; ref. 7). The remainder of the spectrum was essentially the same as that of phosphatidyl choline. Assignments of proton signals and their integration (Table 1) are consistent with the structure of phosphatidyl *S,S*-dimethyl mercaptoethanol. The infrared spectrum of the sulpholipid showed peaks at 2,868 and 2,968 cm^{-1} (CH_2 and CH_3), 1,747 cm^{-1} (ester C = O), 1,475 cm^{-1} (CH_2 and CH_3), 1,285 cm^{-1} (P = O), 1,175 cm^{-1} (ester C–O) and a doublet with a major peak at 1,092 cm^{-1} (PO$^-$) and a shoulder at 1,065 cm^{-1} (P–O–C). There was no evidence of sulphate (1,440–1,350, 1,230–1,150 cm^{-1}) or sulphonate (1,420–1,330, 1,200–1,145 cm^{-1}) bands. Comparison with the spectrum of phosphatidyl choline revealed

Table 2 R_f values and electrophoretic mobilities of the sulphur-containing hydrolysis products from phosphatidyl sulphocholine

Hydrolytic procedure	R_f in butanol–propionic acid–water (142:71:100, v/v)	Relative electrophoretic mobility towards cathode* pH 2.1	pH 6.5
Methanolic HCl†	0.52	1.00	1.00
Mild alkali‡	0.72	1.00	1.00
Phospholipase D	0.52	1.00	1.00
Phospholipase C	0.23	0.00	−0.05
Standards			
Choline	0.57	1.00	1.00
Phosphorylcholine	0.23	0.00	−0.05
2-Hydroxyethyl dimethylsulphonium iodide	0.52	1.00	1.00
2-Methoxyethyl dimethylsulphonium iodide	0.72	0.98	0.97
Trimethylsulphonium iodide	0.54	1.17	1.20
Butyldimethylsulphonium iodide	0.80	0.90	0.89

*Relative to mobility of choline.
†Hydrolysed in 2.5% methanolic HCl for 5 h under reflux; water-soluble phosphates detected were glycerophosphate and methylglycerophosphate (R_f in butanol–propionic acid–water 142:71:100, v/v: 0.17 and 0.30, respectively).
‡According to procedure of Kates[5], neutralised with ethyl formate; the water-soluble phosphate detected was α-glycerophosphate (R_f 0.17 and 0.31 in butanol–propionic acid–water 142:71:100 (v/v) and phenol–water 100:38 (w/w) respectively).

choline and in phosphorylsulphocholine is readily attacked by methoxide ion, resulting in the release of the methoxysulphocholine with concomitant formation of α-glycerophosphate or inorganic phosphate, respectively.

The anomalous behaviour of phosphatidyl sulphocholine towards mild alkali may explain the unusual phospholipid composition of several photosynthetic marine diatoms reported previously[11]. Low amounts of glycerophosphorylcholine (3–8%) and high amounts of α-glycerophosphate (11–32%), as well as a ^{35}S-labelled compound (ref. 11, Fig. 8, spot 11) with mobilities similar to those of methoxysulphocholine, were detected in mild alkaline hydrolysates of the total lipids of these diatoms. Also, a ^{32}P-labelled and ^{35}S-labelled spot with the mobilities of lecithin was detected on chromatograms of the intact lipids of these organisms, indicating the presence of phosphatidyl sulphocholine (ref. 11, Fig. 1 and 7).

From these observations and the findings reported here, it seems that phosphatidyl sulphocholine may be present in high proportions in marine diatoms, replacing phosphatidyl choline as the major membrane phospholipid component. Phosphatidyl sulphocholine would be expected to have a zwitterionic structure at pH 7 and overall size and shape similar to lecithin so that this novel lipid could functionally substitute for lecithin in membranes of N. alba and other diatoms.

The question concerning the biosynthetic pathways for the sulphonium analogue is intriguing, and two possible pathways can be postulated by analogy with those for phosphatidyl choline in mammalian and plant systems[12]

$$\text{sulphocholine} \xrightarrow{\text{ATP,CTP}} \text{CDP-sulphocholine}$$

$$\xrightarrow{\textit{sn}\text{-1,2-diglyceride}} \text{phosphatidyl sulphocholine} \quad (1)$$

$$\beta\text{-mercaptoethanol} \xrightarrow{\text{ATP,CTP}} \text{CDP-}\beta\text{-mercaptoethanol}$$

$$\xrightarrow{\textit{sn}\text{-1,2-diglyceride}} \text{phosphatidyl-}\beta\text{-mercaptoethanol}$$

$$\xrightarrow{2CH_3} \text{phosphatidyl sulphocholine} \quad (2)$$
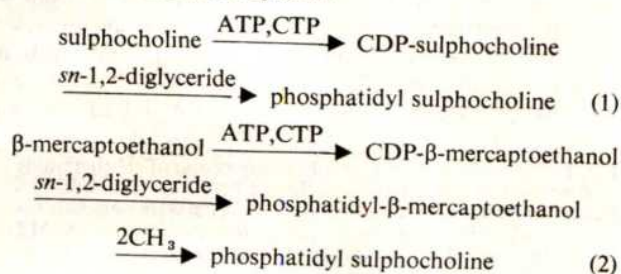
In pathway (1), sulphocholine might be derived from dimethyl β-propriothetin, which is commonly found in marine organisms[13–15] and has been shown to be derived from methionine[14]. Alternatively, the sulphocholine might arise from methylation of β-mercaptoethanol, which is known to be a substrate for methyltransferases in rat liver[16]. Evidence for the operation of pathway (1) has been reported for a mammalian system by Bjerve and Bremer[17] who found that radioactive label from both ^{35}S-sulphocholine and Me-^{3}H-sulphocholine was incorporated in rat heart and kidney phospholipids. With regard to pathway (2), we have not detected the presence of the expected phosphatidyl mercaptoethanol in N. alba, but this would not necessarily eliminate pathway (2) because rapid turnover of the thiol intermediate would make its detection difficult.

The absence in N. alba and other marine diatoms[11] of significant amounts of nitrogenous phospholipids (phosphatidylserine, phosphatidylethanolamine and phosphatidylcholine) and the high concentrations of phosphatidyl sulphocholine and ceramide sulphonate[6] suggest that in these organisms sulphur-containing amino acids (for example, cysteine and methionine) rather than serine, probably act as primary metabolic precursors of the glycerophospholipids and sphingolipids. These novel biosynthetic pathways are being investigated.

R. ANDERSON
M. KATES*

Department of Biochemistry,
University of Ottawa,
Ottawa, Canada K1N 6N5

B. E. VOLCANI

Scripps Institution of Oceanography,
University of California, San Diego,
La Jolla, California 92093

1 Nichols, B. W., Harris, R. V., and James, A. T., Biochem. biophys. Res. Commun., 256–262 (1965).
2 Merdinger, E., J. Bact., 98, 1021–1025 (1969).
3 Brown, A. E., and Elovson, J., Biochemistry, 13, 3476–3482 (1974).
4 Azam, F., and Volcani, B. E., Archs Microbiol., 101, 1–8 (1974).
5 Kates, M., Techniques of Lipidology, (North-Holland, Amsterdam, 1972).
6 Anderson, R., Livermore, B. P., Volcani, B. E., and Kates, M., Biochim. biophys. Acta, 409, 259–263 (1975).
7 Ameja, R., and Chadha, J. S., Biochim. biophys. Acta, 239, 84–91 (1971).
8 Nelson, G. J., Lipids, 3, 104 (1968).
9 Baer, E., Lipids, 3, 384–385 (1968).
10 Shamma, M., Deno, N. C., and Remar, S. F., Tetrahedron. Lett., 13, 1375–1379 (1966).
11 Kates, M., and Volcani, B. E., Biochim. biophys. Acta, 116, 264–278 (1966).
12 Kates, M., and Marshall, M. O., in Recent Advances in Plant Lipid Research (edit. by Galliard, T., and Mercer, E.), 115–159 (Academic, London 1975).
13 Challenger, F., and Simpson, M. I., J. chem. Soc., 1591–1597 (1948).
14 Greene, R. C., J. biol. Chem., 237, 2251–2254 (1962).
15 Motohiro, T., Mem. Fac. Fish. Hokkaido Univ., 10, 1–65 (1962).
16 Bremer, J., and Greenberg, D. M., Biochim. biophys. Acta, 46, 217–226 (1961).
17 Bierve, K. S., and Bremer, J., Biochim. biophys. Acta, 176, 570–583 (1969).

An exosporium-like outer layer in *Bacillus subtilis* spores

BACTERIAL endospores are differentiated cells surrounded by several characteristic spore envelopes—cortex, inner and outer spore coats. In addition to these layers, present in all *Bacillus* and *Clostridium* spores, several species possess a loosely fitting outermost structure—the exosporium[1–5]. In other species, however, including *B. subtilis*, the exosporium has not been found in thin sections[1,6,7], although its possible existence was suggested by the observation of a fibrous 'exosporial' layer by freeze-etching[8]. In our studies of the fine structure of *B. subtilis* spores we have found that a new outer layer, probably an exosporium, is present and can be visualised in thin sections after partial chemical extraction of the spore coats.

PM9, a normally sporulating, auxotrophic (Trp C2, Met C3, Phe A1) strain derived from *B. subtilis* Marburg 168 was used. Spores were obtained in a complex nutrient broth[9] and fixed for electron microscopy either untreated or after incubation for 2 h at 37 °C in 0.1 M Tris, 8 M urea, at pH 8.5 in the presence of 1% mercaptoethanol[10,11]. Fixation, staining and electron microscopy were carried out as described previously[12].

Untreated spores show the classical spore structure with a cortex, an inner coat formed by three well defined layers and a thick, darker outer coat (Fig. 1). After urea–mercaptoethanol treatment (Fig. 2), the cortex and the inner coat remain unchanged, but the outer coat appears more loose and granular. In addition, a loosely fitting single sheet, detached from the

Fig. 1 Untreated free mature spore of *B. subtilis* PM9 exhibiting the typical fine structure of spore envelopes. CX, Spore cortex; IC, inner spore coat; OC, outer spore coat. Note that no exosporium is clearly visible. Bar represents 0.2 μm.
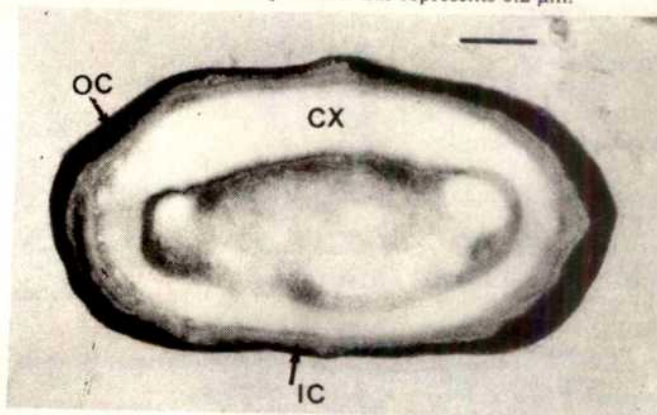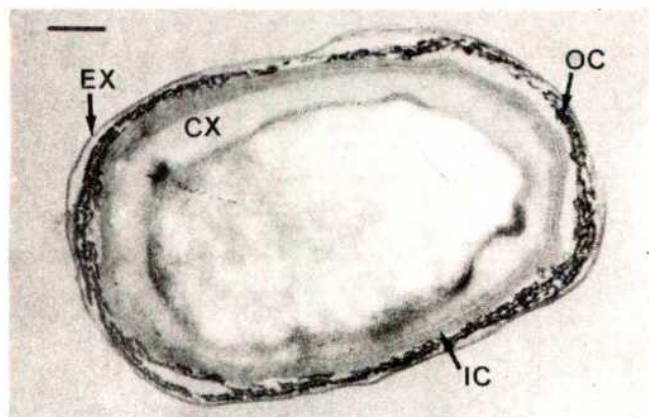
Fig. 2 Free mature spore of *B. subtilis* PM9, treated with urea and mercaptoethanol. The three envelopes visible in the control (Fig. 1) are present. The fine structure of the outer spore coat is disorganised. Note the clearly defined outermost layer which exhibits all the characteristics of an exosporium (EX). Bar represents 0.2 μm.

outer coat, is visible. This layer is very much like the well known exosporium of *B. cereus* and other species[3-5].

Treatment of spores with a reducing agent in the presence of urea is known to liberate a large amount of insoluble, structural proteins from the spore coats[11,12], as reflected in the morphological changes of the coat structure. Our observations suggest the existence of a new, exosporium-like layer, tightly fitting the spore and masked by the dense outer spore coat. In fact, the exosporium is clearly visible in occasional sections of untreated spores, in which it appears tightly fitting the outer surface and remains in most cases undistinguishable from the dark outer coat. The clarity with which the exosporium could be seen, seems to be the result of two effects of the urea–mercapto-ethanol treatment—partial dissolution of the outer coat and loosening of the exosporium.

Similar results were obtained with other wild-type and mutant *B. subtilis* strains. Control experiments show that the treatment which enables visualisation of 'exosporium' does not destroy the viability and heat resistance of the spores. It has been suggested[13] that the exosporium may play a role in the resistance of the spore—that is, against proteases and lytic enzymes. Its presence in *B. subtilis* spores would explain why there is no marked difference between the properties of these spores and those of such species as *B. cereus*.

Demonstration of an exosporium-like structure in *B. subtilis* implies that the classification of bacterial spores into two classes[12], with and without exosporium, is incorrect, and extends the morphological similarity of the spores of various species to the outermost layer.

J. C. F. Sousa
M. T. Silva

Centre of Experimental Cytology,
University of Porto,
Porto, Portugal

G. Balassa

CNRS,
Montpellier, France

1 Murrell, W. G., *Adv. Microbiol. Physiol.*, **1**, 133–251 (1967).
2 Murrell, W. G., Ohye, D. F., and Gordon, R. A., in *Spores*, **IV**, (edit. by Campbell, L. L.), 1–19 (Am. Soc. Microbiol., Bethesda, Maryland, 1969).
3 Gerhardt, P., and Ribi, E., *J. Bact.*, **88**, 1774–1789 (1964).
4 Beaman, T. C., Mankratz, H. S., and Gerhardt, P., *J. Bact.*, **109**, 1198–1209 (1972).
5 Ohye, D. F., and Murrell, W. G., *J. Bact.*, **115**, 1179–1190 (1973).
6 Ryter, A., *Ann. Inst. Pasteur*, **108**, 40–60 (1965).
7 Balassa, G., in *Current Topics in Microbiology and Immunology*, **56**, 99–192 (Springer, Berlin, 1971).
8 Holt, S. C., and Leadbetter E. R., *Bact. Rev.*, **33**, 346–378 (1969).
9 Balassa, G., *Molec. gen. Genet.*, **104**, 73–103 (1969).
10 Gould, G. W., and Hitchins, A., *J. gen. Microbiol.*, 33, 413–424 (1963).
11 Somerville, H. J., Delafield, F. P., and Rittenberg, S. C., *J. Bact.*, **101**, 551–560 (1970).
12 Sousa, J. C. F., Silva, M. T., and Balassa, G., *Molec. gen. Genet.*, **128**, 261–272 (1974).
13 Tipper, D. J., and Gauthier, J. J., in *Spores*, 5, (edit. by Hanson, R., and Campbell, L. L.), 3–12 (Am. Soc. Microbiol. Washington, DC, 1972).

Bacteriophage P22 lysogenises efficiently at high multiplicities of infection because *Salmonella typhimurium* DNA synthetic capacity is limited

Infections of *Salmonella typhimurium* by wild-type phage P22 generally result in a lytic response at low multiplicities of infection (MOIs) and a lysogenic response at high MOIs[1]. At MOIs above 5, more than 90% of all cells are lysogenised. At an MOI of 3, 45% of the infected cells become lysogens whereas at an MOI of 1, only 23% of the infected cells are lysogenised. The regulation of P22 lysogeny has been studied extensively[1-7]. Three structural genes and a *cis*-active site in the immunity C (*immC*) region are required to establish lysogeny. Interaction of gene *c1* and *c3* products at the *c27* site stimulates the production of the *c2* repressor, and at the same time causes a transient retardation in the expression of lytic genes[1,8]. A mutant of *S. typhimurium*, Pox-1, that channels P22 infections at any MOI to lysogeny with an efficiency of 1.0 has a reduced capacity for both host and viral DNA synthesis[6]. Slowed synthesis of DNA in this mutant results in an over-production of regulatory proteins relative to the levels of viral DNA in infected cells. This, in turn, results in lysogeny at all MOIs. We now propose that the initial decision

Fig. 1 Effects of MOI on rates of incorporation of ³H-thymidine. *S. typhimurium* strain 109 was infected at 30 °C with wild-type P22 at the indicated MOIs. Samples (10⁸ cells) were pulse labelled with 20 μCi of ³H-thymidine[5]. ○, Uninfected control; ×, MOI 3; △, MOI 5; ▲, MOI 10. The dotted line is an arbitrary base line indicating possible levels of host DNA synthesis (see text). Dashed arrows mark the areas between 6 and 15 min after infection.
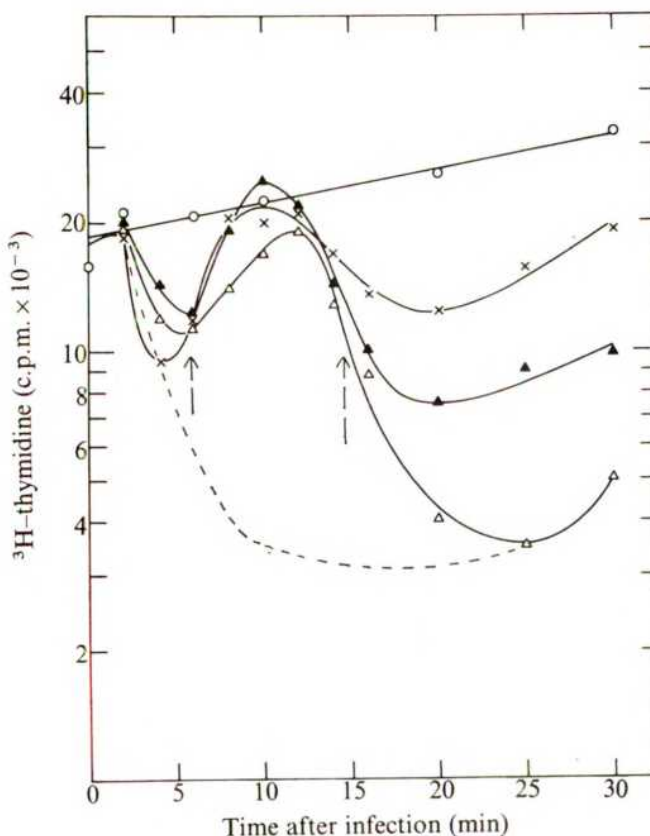
Table 1 Percentage hybridisation of DNA and RNA from infected cells to P22 DNA

	MOI				
	1	2	3	10	20
(a) DNA–DNA hybridisation	24%	19%	21%	44%	32%
(b) RNA–DNA hybridisation	—	—	1.8%	6.3%	10.3

a, Hybridisation of ^{3}H-thymidine-labelled DNA from infected cells to P22 DNA. ^{3}H-thymidine (200 µCi ml[−1]) was added to strain 109 6 min after infection and the cells were lysed 20 min after infection as described before[11]. Hybridisation to P22 DNA was done by the method of Denhardt[12]. Data are expressed as the percentage of input counts bound to P22 DNA. Values for MOIs of 1 and 2 were corrected for the presence of uninfected cells in those cultures. *b*, Hybridisation of ^{3}H-uridine-labelled RNA from infected cells to P22 DNA. RNA was pulse labelled with ^{3}H-uridine (15 Ci ml[−1]) for 1 min, 8 min after infection of strain 109 by wild-type P22. RNA was prepared and hybridised to P22 DNA by the method of Smith[13].

between lysis and lysogeny in wild-type cells is also a function of the relative amounts of phage DNA and the phage proteins from the *immC* region.

Wild-type cells were infected with wild-type P22 at MOIs of 3, 5, and 10, and the rates of ^{3}H-thymidine incorporation were determined (Fig. 1). The initial decline in incorporation rate during the first 6 min was a result of shutoff of host DNA synthesis. The extent of host shutoff was comparable in the three infections. Phage DNA synthesis is required for both lysis and lysogeny. The areas under the curves between 6 and 15 min represent such synthesis[4]. Incorporation of ^{3}H-thymidine between 15 and 30 min at MOIs 3 and 5 was primarily into cells which were lytically infected. An arbitrary base line of host DNA synthesis rate is drawn at 3×10^3 c.p.m. min[−1]. This is the level seen by Levine and Schott[4], using a phage mutant unable to synthesise DNA. The areas under the curves between 6 and 15 min were integrated to estimate the relative amounts of viral DNA synthesised in the three infections. Total calculated c.p.m. are 9.1×10^4, 7.1×10^4, and 9.0×10^4 at MOIs of 3, 5 and 10, respectively. The ratios of total c.p.m. incorporated (1.29 : 1.0 : 1.28) are different from the ratios of the parental genomes available for replication. These ratios, and the fact that the heights of the three peaks are comparable, suggest that the maximum replication of parental genomes was reached at an MOI of 3 or less.

Table 1*a* shows the results of hybridising DNA from infected cells to P22 DNA attached to nitrocellulose filters. The ratio of percentage hybridisation for MOI 1 to that for MOI 20 is 1.0 : 1.33, and is much different from the ratio of the parental genomes. It is reasonable to suggest that near maximum replication of P22 DNA occurs at an MOI of 1. Moreover, these data support our conclusion that host shutoff does not depend on MOI, for the fraction of DNA that is virus specific is nearly constant.

RNA synthesis did not follow the same pattern as DNA synthesis. Figure 2 shows the incorporation of ^{3}H-uridine pulses into acid-precipitable material. There were almost fourfold differences in both initial decreases and in maximum levels of incorporation seen at MOIs of 3 and 20. Differences in the degree to which host specific transcription was shutoff probably accounts for the variation in incorporation during the first 2–4 min. This makes interpretation of the curves difficult, since each curve represents the sum of host and phage-specific RNA, and the amount of host RNA may be different in each infection.

We determined the amount of viral-specific RNA synthesised 8 min after infection by hybridising RNA from infected cells to P22 DNA. The results are shown in Table 1*b*. As MOI increased, the total incorporation of ^{3}H-uridine decreased (Fig. 2), but the fraction that was virus specific increased. Two conclusions are drawn from these data. First, the synthesis of phage RNA was directly dependent
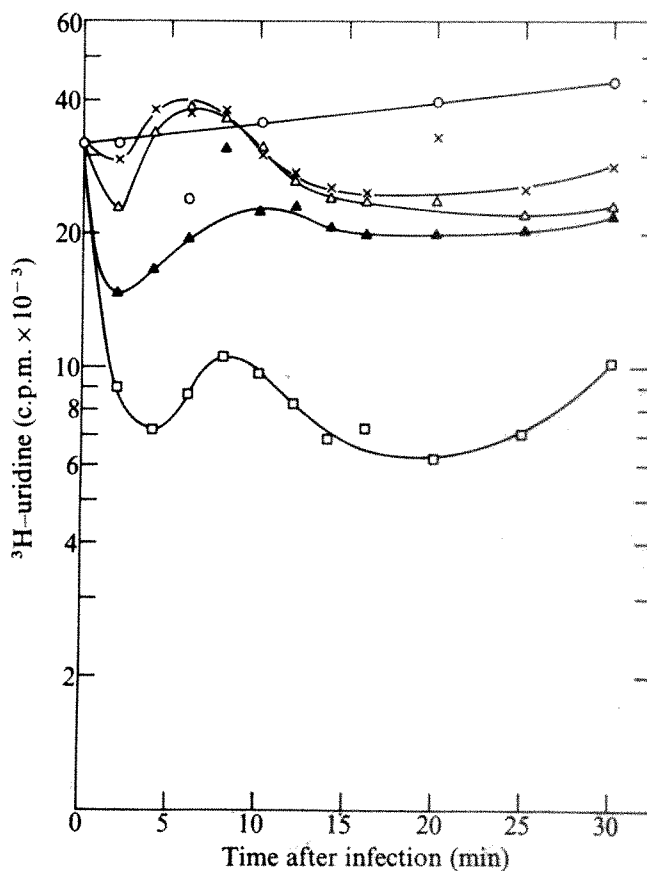
Fig. 2 Effects of MOI on rates of incorporation of ^{3}H-uridine. Strain 109 was pulse labelled with 40 µCi of ^{3}H-uridine using the same procedure as for ^{3}H-thymidine incorporation[5]. ○, Uninfected control; ×, MOI 3; △, MOI 5; ▲, MOI 10; □, MOI 20.

on MOI. The ratio of the percentage hybridisable counts at MOI 3 to percentage hybridisable counts at MOI 20 is 1.0 : 5.5, close to the ratio of MOI 3 to MOI 20 which is 1.0 : 6.6. Second, because the fraction of RNA that was virus specific increased with MOI, the fraction that was host specific must have decreased. Therefore, the extent of host shutoff after infection was also MOI dependent.

There is no evidence for translational controls in P22 infections. Therefore, increased levels of RNA synthesis at high MOIs should result in increased levels of protein synthesis. We expect the *cl* gene transcript to be among the few species of phage RNA that are made 8 min after infection[1]. The *cl* protein is pivotal in the establishment of lysogeny[1−3,5−8]. If higher MOIs result in enhanced transcription of this gene, then the ratio of *cl* protein to DNA would be increased at higher MOIs. At present, there is no assay for the *cl* protein. Therefore we are restricted to measurements of RNA.

If cells are infected with wild-type P22 at various MOIs, and the percentage of infected cells receiving 3 or more copies of the parental genome is calculated from a Poisson distribution, there is good agreement between these values and the percentage of infected cells that become lysogens. We suggest that in a cell infected with 1 or 2 viral genomes, the rate of DNA synthesis relative to RNA synthesis is sufficiently rapid that not all the viral DNA can be complexed with *cl* protein, and the infection proceeds to lysis. If three or more copies of phage DNA enter a host cell, the rate of production of *cl* protein is sufficient to bind all the DNA and the cell is lysogenised. The dependence of lysogeny on MOI is explained by the limited capacity of the infected cell to make DNA; this is saturated at MOIs near 1. RNA synthesis is not so limited. Therefore, at high MOI,

the ratio of RNA synthesis (and regulatory proteins) to DNA synthesis is enhanced. When the ratio of regulatory proteins to DNA is sufficiently high, infections go to lysogeny. This is achieved at MOIs of 3 and more in P22.

Phage P22 has been classified as a lambdoid phage[9]. Phage λ also shows increased frequencies of lysogeny at higher MOI[10]. We expect that limited DNA synthetic capacity of the infected bacteria accounts for the increased frequencies.

This research was supported by grants from the NSF and NCI.

BETTIE M. STEINBERG
MICHAEL GOUGH

Department of Microbiology,
State University of New York,
Stony Brook, New York 11794

1 Levine, M., *Virology.*, 3, 22–41 (1957).
2 Gough, M., and Tokuno, S., *Molec. gen. Genet.*, 138, 71–79 (1975).
3 Levine, M., and Smith, H., *Science*, 146, 1581–1582 (1964).
4 Levine, M., and Schott, C., *J. molec. Biol.*, 62, 53–64 (1971).
5 Smith, H., and Levine, M., *Proc. natn. Acad. Sci. U.S.A.*, 52, 356–363 (1964).
6 Steinberg, B. M., and Gough, M., *J. Virol.*, 16, 1154–1160 (1975).
7 Tokuno, S., and Gough, M., *J. Virol.*, 16, 1184–1190 (1975).
8 Tokuno, S., and Gough, M., *Molec. gen. Genet.*, 144, 199–204 (1976).
9 Botstein, D., and Herskowitz, I., *Nature*, 251, 584–549 (1974).
10 Fry, B. A., *J. gen. Microbiol.*, 21, 676–684 (1959).
11 Gough, M., Scott, J. V., Malik, V. S., and De La Rosa, O., *Virology*, 47, 276–284 (1972).
12 Denhardt, D. T., *Biochem. biophys. Res. Commun.*, 23, 641–646 (1966).
13 Smith, H. O., *Virology*, 34, 203–223 (1968).

Acquired resistance to infection with *Schistosoma mansoni* induced by *Toxoplasma gondii*

SPECIFIC acquired immunity to *Schistosoma mansoni* in the mouse has only recently been demonstrated definitively[1]. This phenomenon does not occur until 12 weeks after primary infection and has been found to be adoptively transferred only by serum[2]. The induction of so-called 'immunity' to *S. mansoni* has also been reported in mice and hamsters exposed to a wide variety of unrelated organisms and extracts[3-6], but in most, if not all cases it is likely to be nonspecific. Although all of these acquired forms of resistance have been lumped together under the name of immunity, it is probable that the mechanisms of specific and nonspecific resistance are totally different. The present investigation of the effect on *S. mansoni* infection of *Toxoplasma gondii*, an intracellular protozoan parasite which renders mice resistant to a variety of unrelated organisms[7-11], largely by activation of macrophages[12], may clarify the role of nonspecific mechanisms in protection against schistosomiasis.

Swiss albino female mice (18–20 g) were purchased from Carworth Farms, New City, New York, although in one experiment mice were obtained from Flow Laboratories, Rockville, Maryland. *T. gondii* cysts were obtained from the brains of animals infected for 5–6 weeks with the Gleadle strain. Experimental animals were injected intraperitoneally with 10 cysts suspended in 0.2 ml brain emulsion to which penicillin (100 U ml^{-1}) and streptomycin (10 μg ml^{-1}) were added. Similar volumes of brain emulsion from uninfected mice containing antibiotics were injected intraperitoneally into control animals.

Recovery of immature schistosomes (schistosomula) from the lungs was carried out in two separate experiments in groups of eight *Toxoplasma*-infected mice and eight control animals (injected with normal mouse brain emulsion) 2 weeks after the initiation of the protozoan infection. The animals were anaesthetised and exposed percutaneously to 500 cercariae of a Puerto Rican strain of *S. mansoni*[13]. Recovery of schistosomula was carried out on day 6 after infection (the day of maximum recovery for the strain of mice used) as described by Sher *et al.*[1].

For quantitation of adult worms, mice were injected subcutaneously with 20 cercariae of a Puerto Rican strain of *S. mansoni*. Six to eight weeks after infection, groups of 10 animals were anaesthetised and the portomesenteric venous system perfused[14]. In addition, the livers and mesenteric vessels were searched for worms. All worms recovered from each experimental group were washed in normal saline, counted, and the length of the paired worms measured by an ocular micrometer mounted on a dissecting microscope.

In one experiment the worms were washed three times in normal saline, emulsified in a tissue grinder and injected intraperitoneally into four mice to determine whether or not the schistosomes harboured any *T. gondii*. The animals were killed 6 weeks later, their brains examined for cysts, and their sera tested for fluorescent antibodies[15].

Schistosome eggs were counted by digesting the livers removed from *Toxoplasma*-infected and control mice 8 weeks after infection with *S. mansoni*[16].

In two experiments using mice, respectively, from Carworth Farms and Flow Laboratories, the *Toxoplasma*-infected animals had 35% (61 ± 5 compared with 94 ± 8 in the controls) and 40% (89 ± 10 compared with 148 ± 12 in the controls) reductions in the number of schistosomula recovered 6 d after exposure to *S. mansoni* cercariae. The differences between the experimental and control animals in each case were highly significant ($P < 0.01$).

Table 1 summarises the results of adult worm perfusions 6–8 weeks after *S. mansoni* infection. In two separate experiments, animals infected with *T. gondii* 1 d before exposure to *S. mansoni* developed 32% and 22% fewer total worms, and in each case 50% fewer worm pairs in comparison with controls. The differences in each experiment were significant at the 5% level. Furthermore, measurement

Table 1 Effect of injection of ten *T. gondii* cysts at different times before and after exposure to 20 *S. mansoni* cercariae: numbers and sizes of the adult worms 8 weeks after induction of schistosomiasis

	Mean number of worms ± s.e.				Mean length of worms (mm ± s.e.)	
	Males	Females	Pairs	Total	Males	Females
T. gondii 4 weeks before:						
Controls	0.8 ± 0.2	1.0 ± 0.4	3.4 ± 0.5	8.5 ± 0.8	9.2 ± 0.4	12.1 ± 0.4
T. gondii-infected	1.1 ± 0.4	0.6 ± 0.4	2.0 ± 0.4	5.5 ± 0.7	9.0 ± 0.2	11.8 ± 0.6
T. gondii 1 d before:						
Experiment 1						
Controls	3.0 ± 0.5	1.4 ± 0.5	3.2 ± 0.4	10.8 ± 1.2	8.9 ± 0.2	10.8 ± 0.3
T. gondii-infected	2.9 ± 0.5	0.9 ± 0.4	1.6 ± 0.3	7.3 ± 0.8	8.4 ± 0.2	8.8 ± 0.4
Experiment 2						
Controls	2.5 ± 0.6	1.0 ± 0.3	2.4 ± 0.6	9.3 ± 1.1	8.0 ± 0.4	9.6 ± 0.4
T. gondii-infected	3.0 ± 0.5	1.1 ± 0.3	1.2 ± 0.2	6.5 ± 0.4	6.5 ± 0.2	7.4 ± 0.2
T. gondii 4 weeks after:						
Controls	1.1 ± 0.3	0.2 ± 0.1	3.3 ± 0.2	7.9 ± 0.4	8.8 ± 0.4	11.1 ± 0.5
T. gondii-infected	1.1 ± 0.4	0	3.4 ± 0.3	8.5 ± 0.4	9.0 ± 0.5	10.9 ± 0.6

57

of the paired worms in both experiments showed that in each case the female worms from the *Toxoplasma*-infected mice were significantly shorter ($P<0.05$) (Table 1). The effect of the time of induction of toxoplasmosis on the recovery of adult schistosomes was then studied. *Toxoplasma* cysts were injected 4 weeks before and 4 weeks after the initiation of the schistosome infection (Table 1). Whereas previous toxoplasmosis resulted in a 35% reduction in the total worm load, infection 4 weeks after the injection of *S. mansoni* cercariae had no effect.

There was no evidence that the adult schistosomes were infected with *T. gondii* as the four mice injected 6 weeks previously with emulsified worms from *Toxoplasma*-infected animals had no cysts in their brains and no antibodies in their sera.

The results of *S. mansoni* egg counts in the livers of the different experimental groups were as follows: control animals injected with normal brain emulsion developed a mean egg load per liver of $6,736 \pm 325$, whereas animals with *T. gondii* infection induced 4 weeks or 1 d before schistosomiasis had respective reductions in the egg counts of 35% ($4,480 \pm 210$) and 43% ($3,840 \pm 301$). In contrast, animals infected with *T. gondii* 4 weeks after *S. mansoni* had mean egg counts similar to the controls ($6,077 \pm 286$).

Toxoplasmosis has been reported to protect mice against a variety of unrelated organisms and tumours[7-11]. This effect has been ascribed largely to activation of the macrophages[12]. Ruskin *et al.*[12] have demonstrated that activated macrophages from *T. gondii*-infected mice not only display activity *in vivo* but have the capacity *in vitro* to destroy nonspecific target cells by a mechanism which does not involve phagocytosis. It has also been shown that other intracellular infectious agents such as *Besnoitia jellisoni* confer similar properties on mouse peritoneal macrophages[11].

Schistosomiasis is one of the major helminth infections of mankind, and greater understanding of the host defence mechanisms is urgently needed. The variety of conditions discussed under the blanket heading of immunity, however, tends to be confusing. Only recently has it been demonstrated definitively that specific acquired resistance to re-infection occurs in mice previously infected with *S. mansoni*. This does not occur until 12 weeks after infection and has been transferred passively with serum but not with cells[1,2]. Using monospecific antieosinophil serum[17] it has been shown that this form of resistance is due to an anti-body-dependent cell-mediated mechanism in which the effector cell is the eosinophil[18].

Also considered under the term immunity have been various forms of resistance induced by snail haemolymph[3], monkey erythrocytes[4], formalin-killed *Escherichia coli*[5] and extracts of *Fasciola hepatica*[6]. It seems probable, however, that most, if not all of these materials induced a nonspecific form of acquired resistance involving mechanisms totally different from that of specific immunity. The present experiments in which *T. gondii* induced resistance in mice to *S. mansoni* strongly suggest that nonspecific resistance can be acquired in schistosomiasis.

Although *T. gondii* activates macrophages[12], a factor considered to be largely responsible for the nonspecific protection it affords against other organisms, this protozoan is also a powerful suppressant of B and T lymphocytes[14,20], cells which are necessary for the induction and maintenance of specific immunity. Furthermore, specific acquired resistance in the mouse takes about 3 months to develop, but the protective effects of *T. gondii* and the other nonspecific agents mentioned above are fully manifest within a few days.

With respect to our knowledge of the immunology of schistosomiasis, therefore, it is important not to lump all forms of acquired resistance to infection under the general term immunity, but to determine whether they are specific

or nonspecific. Since both types provide only partial resistance, and their kinetics and mechanisms are different, perhaps their effects may be additive.

ADEL A. F. MAHMOUD
KENNETH S. WARREN
G. THOMAS STRICKLAND

Division of Geographic Medicine,
Department of Medicine,
Case Western Reserve University
and University Hospitals,
Cleveland, Ohio 44106
and Division of Immunoparasitology,
Naval Medical Research Institute,
Bethesda, Maryland 20014

1 Sher, A., Mackenzie, P., and Smithers, S. R., *J. infect. Dis.*, 130, 626–633 (1974).
2 Sher, A., Smithers, R. S., and Mackenzie, P., *Parasitology*, 70, 347–357 (1975).
3 Oliver-Gonzales, J., *Proc. Soc. exp. Biol. Med.*, 128, 1029–1033 (1968).
4 Oliver-Gonzales, J., *Am. J. trop. Med. Hyg.*, 16, 565–567 (1967).
5 Smith, M. A., Clegg, J. A., Kusel, J. R., and Webbe, G., *Experientia*, 31, 595–596 (1974).
6 Hillyer, G. V., de Diaz, A. L., and Garcia-Blanco, M., *Bol. Asoc. Med., PR*, 67, 381 (1975).
7 Ruskin, J., and Remington, J. S., *Science*, 160, 72–74 (1968).
8 Remington, J. S., and Merigan, T. C., *Proc. Soc. exp. Biol. Med.*, 131, 1184–1188 (1969).
9 Gentry, L. O., and Remington, J. S., *J. infect. Dis.*, 123, 22–31 (1971).
10 Hibbs, J. B., Lambert, L. H., and Remington, J. S., *J. infect. Dis.*, 124, 587–592 (1971).
11 Ruskin, J., and Remington, J. S., *Antimicrob. Agents Chemother.*, 474–477 (1968).
12 Ruskin, J., McIntosh, J., and Remington, J. S., *J. Immun.*, 103, 252–259 (1969).
13 Warren, K. S., and Peters, P. A., *Am. J. trop. Med. Hyg.*, 16, 718–722 (1967).
14 Duvall, R. H., and DeWitt, W. B., *Am. J. trop. Med. Hyg.*, 16, 483–486 (1967).
15 Anderson, S. E., and Remington, J. S., *South. Med. J.*, 68, 1433–1443 (1975).
16 Cheever, A. W., *Bull. WHO*, 39, 328–331 (1968).
17 Mahmoud, A. A. F., Warren, K. S., and Boros, D. L., *J. exp. Med.*, 137, 1526–1531 (1973).
18 Mahmoud, A. A. F., Warren, K. S., and Peters, P. A., *J. exp. Med.*, 142, 805–813 (1975).
19 Huldt, G., Gard, S., and Olovson, S. G., *Nature*, 244, 301–303 (1973).
20 Strickland, G. T., Ahmed, A., and Sell, K. W., *Clin. exp. Immun.*, 22, 167–176 (1975).

Malaria transmission blocked by immunisation with gametes of the malaria parasite

IN view of the limited success of malaria eradication schemes based on concepts of vector control and chemotherapy, serious efforts are being made to investigate the possibility of developing an antimalarial vaccine. So far two classes of vaccine have been under investigation: (1) an anti-sporozoite vaccine designed to prevent malarial infection by blocking the infectivity of sporozoites introduced by mosquito bite[1,2], and (2) vaccines directed against the asexual stages in the blood[3-5]. Although the potential value of such vaccines is unquestionable, their realisation faces technical problems[6]. The latest proposition is to induce immunity against the parasite stages which infect the mosquito and by so doing prevent transmission of the disease by the vector. Gwadz has shown that the infectivity of malarious chickens to mosquitoes can be reduced greatly by prior vaccination with formalin-treated or X-irradiated blood infected with the malaria parasite *P. gallinaceum*[7]. The number of oocysts (the stage of the parasite developing on the wall of the mosquito gut) in mosquitoes fed on vaccinated chickens was reduced by 95–98% below that recorded in mosquitoes fed unvaccinated birds. To achieve this reduction, chickens were given three weekly intravenous inoculations of a total of 4.5 ml of formalin-treated or irradiated blood. Using the same schedule, but with partially purified gametes of the malaria parasite, we have now reduced the infectivity of malarious chickens to mosquitoes at least 99.9% below control levels. To achieve
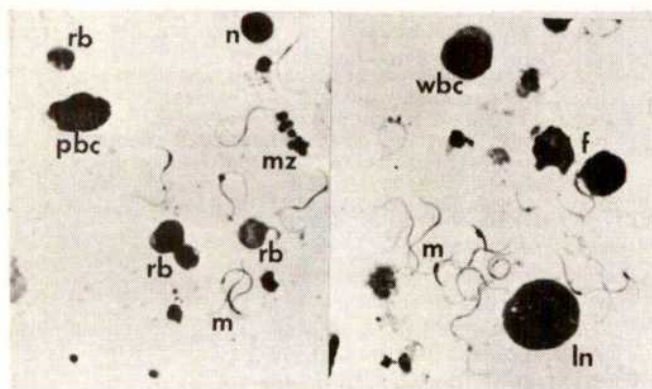
Fig. 1 The material in the 18,000g pellet (see text) used for vaccinating chickens. The preparation has been fixed and stained with Geimsa for the purposes of illustration. m, Male gamete; f, female gamete; rb, residual body of male gametocyte (precursor cell of male gametes) after release of gametes; pbc, parasitised red blood cell; mz, merozoites (mature asexual cells ready to invade red blood cells); n, nucleus of red blood cell within red cell ghost membrane; ln, lysed nucleus of red blood cell; wbc, white blood cell.

this reduction we had to use a minimum of 0.1 μl of packed cellular material.

We used free gametes for the following reasons. To establish a malarial infection in the mosquito after a blood meal, male and female gametes of the parasite must undergo fertilisation in the lumen of the mosquito midgut. We predicted that this usually rapid process might be blocked by antibodies directed against one or both types of gamete. In the ordinary course of events the vertebrate host is apparently never exposed to the free gametes, which are released in malarious blood only in the gut of the mosquito or on exposure to the atmosphere. It seemed reasonable to assume, therefore, that in natural conditions transmission proceeds in the absence of significant immunity against these stages in the life cycle of the parasite. By vaccinating with and artificially inducing immunity against the free gamete we hoped to block fertilisation and hence prevent transmission of the malaria parasite to the mosquito.

Blood from chickens with high (40–60%) but increasing parasitaemias of *P. gallinaceum* was drawn from the heart and washed immediately in 50 volumes of suspended animation (SA) medium (0.21% Tris; 0.96% NaCl; 0.2% glucose, adjusted to pH 7.4). Suspension of blood in SA medium facilitates removal of factors interfering with fertilisation and transmission, which may be present in the plasma, while the release of malarial gametes (usually spontaneous on exposure of blood to air) is reversibly suppressed. The

suspension of blood in SA medium was spun at 500g for 5 min and the packed cells were resuspended in a medium that induced the release of the gametes from their precusor cells in the blood (100 volumes of inactivated foetal bovine serum; 20 volumes of 1.46% NaHCO₃; 2.5 volumes of 5% NaCl, and 2.5 volumes of 10% glucose, adjusted to pH 8.0). The cells were incubated for 30 min at room temperature in 2 volumes of the gamete-releasing medium for each volume of packed cells to facilitate complete release of both male and female gametes, the suspension being readjusted to pH 8.0. After centrifugation of the cell suspension for 5 min at 500g the supernatant contained numerous gametes (male and female), some extracellular asexual parasites, red cells, white cells and acellular debris. The supernatant was then spun at 18,000g for 10 min. By visual estimation (comparison of the pellet with measured volumes of methylene blue in identical centrifuge tubes) the amount of material in the 18,000g pellet occupied between one and two thousandths of the packed cell volume of the blood from which it was derived. This was the material used for vaccination (Fig. 1). In view of the heterogeneous nature of the cell and subcellular material in this preparation it was difficult to quantify accurately the relative amounts of gamete and non-gamete material. The yield of male gametes was about 10% of that in the original blood. Male gametes represented about 50% and female gametes about 20% of the cells present. A detailed account of the method of collection of gametes will be published elsewhere.

The 18,000g pellet was resuspended in SA medium and X-irradiated with 20,000 rad. The concentration of male gametes in the suspension was counted using a haemocytometer and the material was diluted in SA medium by threefold steps to give a series of six dilutions (Table 1). Six, 3-week-old, New Hampshire Red chickens were inoculated intravenously with three injections of the irradiated gametes (freshly prepared) in 1.5 ml of SA medium at weekly intervals. The total number of male gametes received by each chicken varied from 9.0×10^4 to 2.2×10^7, representing the yield from 0.05 ml to 11.0 ml of blood at a 40–60% parasitaemia (Table 1). One week after the last inoculation of gametes the vaccinated chickens and one unvaccinated chicken of the same age were challenged with *P. gallinaceum*-infected blood. Peak parasitaemias of 40–60% were achieved in all vaccinated chickens compared with a peak parasitaemia of 58% in the control. A fresh batch of *Aedes aegypti* mosquitoes was fed daily on each chicken throughout its period of patent parasitaemia (defined here as 0.2% or more of blood cells parasitised). Seven days after feeding, 10 mosquitoes from each batch were dissected and their guts were examined for oocysts. The course of patent parasitaemia and infectivity to mosquitoes in experimental and control chickens are shown in Fig. 2. The total number of oocysts produced and the mean number of

Table 1 Oocyst production in mosquitoes fed on vaccinated and unvaccinated chickens

Total immunising dose of male gametes	Volume of parasitised blood from which the gametes were prepared (ml)	Mean no. of oocysts per mosquito gut during the period of patent blood infection	Total no. of oocysts produced throughout the course of infection by feeding 10 mosquitoes per d
Unvaccinated	—	25.2	2,267
9.0×10^4	0.05	8.2	411
2.7×10^5	0.16	0.00	0
8.3×10^5	0.47	0.01	1
2.5×10^6	1.4	0.00	0
7.4×10^6	3.8	0.01	1
2.2×10^7	11.0	0.00	0

Each chicken was vaccinated intravenously at three weekly intervals, the same amount of X-irradiated gamete containing material being inoculated each week. The total amount received by each chicken is indicated in terms of the number of male gametes in the three inocula and the volume of parasitised blood from which the preparations were derived. A single unvaccinated chicken was used as a control. After challenge with live parasites a fresh batch of mosquitoes was fed daily on each chicken during the course of patent blood parasitaemia (>0.2% of blood cells parasitised). Seven days after feeding ten mosquitoes from each batch were dissected and the oocysts on each mosquito midgut were counted.

Fig. 2 The course of patent blood parasitaemia ($>0.2\%$ of blood cells parasitised) and infectivity to mosquitoes (mean number of oocysts per gut in mosquitoes fed during patent parasitaemia) in vaccinated and unvaccinated chickens. $+$, Unvaccinated chickens; ▲, chicken vaccinated with 9.0×10^4 male gametes; △, 2.7×10^5; ●, 8.3×10^5; ■, 2.5×10^6; ○, 7.4×10^6; □, 2.2×10^7.
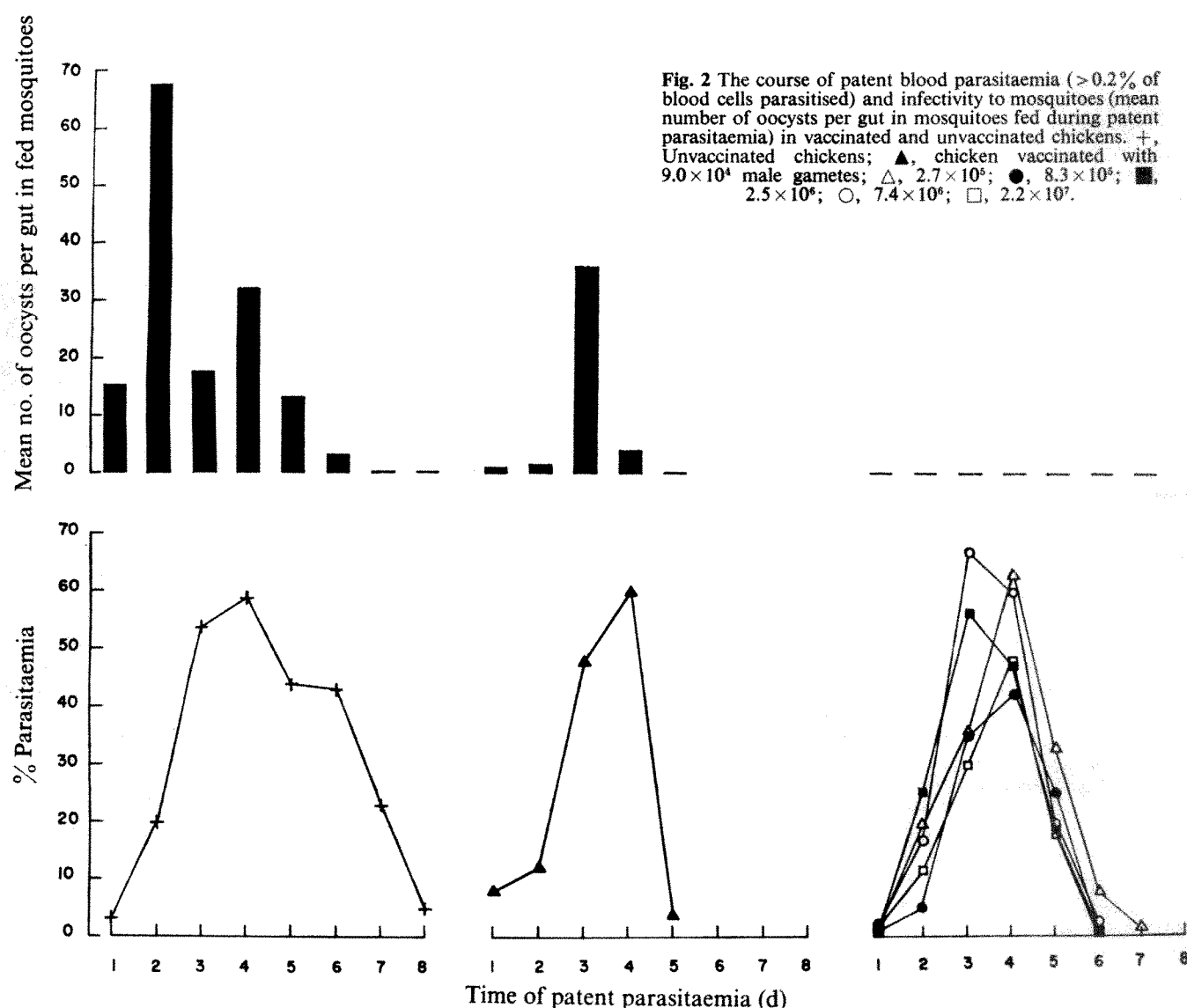
oocysts per mosquito over the total period of patent infection in each chicken is shown in Table 1.

In a normal infection of *P. gallinaceum* in the chicken the total number of oocysts produced by feeding 10 *A. aegypti* per day throughout the course of infection lies in the range of 2,000–4,000. In our experiment the mosquitoes fed on the unvaccinated chicken yielded more than 2,000 oocysts, representing an average of about 25 oocysts per mosquito during the period of transmission by mosquitoes. By contrast, all but the chicken vaccinated with the fewest gametes were almost totally uninfective to mosquitoes throughout their period of patent infection, oocyst counts being reduced by 99.9–100%. Even for the chicken vaccinated with the fewest number of gametes (9×10^4 gametes equivalent to the yield of 0.05 ml of blood) infectivity, as measured by oocyst production, was reduced by 80% below that of the control. Similar results have been obtained using comparatively highly purified preparations of female gametes as a vaccine and with highly purified male gametes.

Oocyst counts in infected anophelene mosquitoes in areas of holoendemic human malaria are generally less than 10 per gut[8]. In principle the level of suppression of oocyst production achieved in our study should lead to considerable reductions in the infectivity of vectors of human malaria if the same response to gamete immunisation could be reproduced successfully in exposed human populations.

We have demonstrated that a transmission-blocking vaccine based on malarial gametes is easily prepared, uncomplicated in administration (effective without adjuvants) and highly potent in the *P. gallinaceum*–chicken system. But as in the case of other forms of antimalarial vaccines tested with animal systems the translation of a gamete-based vaccine into a realistic measure for combating human malaria faces formidable problems. Not only must the effectiveness of a gamete vaccine be demonstrated for human malarias, but the means of producing large quantities of suitable material must be found. Nevertheless, if satisfactory progress is made in these areas we believe that a vaccine based on the free gametes of the malaria parasite may have considerable significance for the prospects of malaria control or eradication.

The concepts arising from this study may have relevance beyond their application to malaria. In most vector-borne protozoal or helminthic diseases (for example, malaria, babesiosis, trypanosomiasis, lieshmaniasis and filariasis) the disease organisms undergo marked biological transformations after their entry into the vector. It is possible that in some instances vaccination of the host with parasite material from the earliest stages of transformation in the vector may reduce transmission of the disease.

for criticism and suggestions in the preparation of the manuscript.

RICHARD CARTER
DAVID H. CHEN

Malaria Section,
Laboratory of Parasitic Diseases,
National Institute of Allergy and Infectious Diseases,
National Institutes of Health,
Bethesda, Maryland 20014

1 Nussenzweig, R. S., Vanderberg, J. P., Most, H., and Orton, C., *Nature*, **216**, 160–162 (1967).
2 Clyde, D. F., Most, H., McCarthy, V. C., and Vanderberg, J. P., *Am. J. med. Sci.*, **266**, 169–177 (1973).
3 Brown, K. N., Brown, I. N., and Hills, L. A., *Expl Parasit.*, **28**, 304–317 (1970).
4 Simpson, G. L., Schenkel, R. H., and Silverman, P. H., *Nature*, **247**, 304–305 (1974).
5 Mitchell, G. H., Butcher, G. A., and Cohen, S., *Immunology*, **29**, 397–407 (1975).
6 Miller, L. H., *Wld Hlth Org. tech. Rep. Ser. No. 579 Expert Committee on Immunization against Malaria* (1975).
7 Gwadz, R. W., *Science* (in the press).
8 Muirhead-Thomson, R. C., *Trans. R. Soc. Trop. Med. Hyg.*, **48**, 208–225 (1954).

Induction of autoantibodies to red blood cells by polyclonal B-cell activators

THE ability to distinguish self from non-self is a fundamental property of the lymphoid system. One hypothesis for immunocyte triggering[1] proposes that the encounter between antigen and immunoglobulin receptors results in irreversible tolerance, whereas activation requires additional signals, thus ensuring that tolerance to self-antigens will always be the first and easiest event to achieve. An alternative view is that the self–non-self discrimination is carried out by T cells[2,3], whereas B cells could never be made tolerant to thymus-dependent antigens, including autologous antigens, because these antigens cannot deliver any signal to B cells by interacting with the Ig or any other B-cell receptor[4,5]. Experimental support for this notion has been obtained by the demonstration that tolerance to autologous antigens can be broken by the injection of cross-reacting antigens[6] and that there are antigen-binding human B cells able to bind human thyroglobulin[3]. To distinguish between the two hypotheses, it is necessary to determine whether tolerance to self-antigens exists at the B-cell level. This can be done by treating lymphocytes with polyclonal B-cell activators (PBA) and studying antibody formation against autoantigens, for PBA are competent to active B cells of any specificity and thus to reveal the total V gene repertoire of the B cells. We have activated bovine spleen cells in culture with PBA (lipopolysaccharide (LPS) from *Escherichia coli* 055 : B5 and purified protein derivative (PPD) of tuberculin) and studied the number of plaque-forming cells (PFC) appearing against autologous red cells and heterologous red cells coated with the hapten NNP. In all experiments, the spleen cells and the red cells were derived from the same animal. We found that PBA-activated bovine lymphocytes synthesised autoantibodies agains their own red cells.

As shown in Fig. 1, which represents the results from three cows, there was a response measured as thymidine incorporation to different concentrations of LPS, although to a lesser extent than with mouse spleen cells. But LPS induced activation, as determined by an increased thymidine incorporation, was highly significant, being three to six times the background values.

To study induction of autoantibody synthesis, bovine spleen cells were cultured in serum-free medium and stimulated with different concentrations of LPS and in some experiments also with PPD. After 2 d PFC against autologous red cells and hapten (NNP)-coated sheep red cells were counted. Bovine lymphocytes were used because bovine red cells can be used in the plaque assay[7]. As Fig. 2 and

Table 1 show, antibodies were always induced to NNP-coated sheep red cells although there were fewer PFC than expected from mouse experiments. There was, however, an analogous increase of the number of PFC against untreated autologous red cells. There were usually fewer PFC against autologous cells than against NNP–SRC, but the two showed the same increase over background and the same dose–response curve.

Thus, the polyclonal B-cell activators LPS and PPD could both induce the formation of autoantibodies by bovine spleen cells. These substances also induced thymidine uptake and plaque formation against NNP-coated sheep red cells.

It seemed possible that LPS and PPD would coat the red cells and that the PFC actually detected activity against these added foreign materials. But this possibility can be excluded on the basis of previous experiments[8], because the concentrations of PBA used here were such that antibody formation against the PBA was completely inactivated, whereas polyclonal antibody synthesis was induced. Thus there must be resting B lymphocytes in bovine spleen cells that can be activated, by the addition of PBA, to synthesise antibodies able to lyse their own red cells.

These findings have far reaching implications for an understanding of the mechanism of self–non-self discrimination. They indicate that B cells are not normally tolerant to their own thymus-dependent self constituents. This agrees with findings on experimentally induced tolerance to a thymus-dependent protein antigen in mice, where completely tolerant animals had a normal complement of B lymphocytes able to produce antibodies against the tolerogen after activation by LPS (refs 9, 10).

The findings also have implications for an understanding of the mechanism of B-lymphocyte activation. According to the two-signal concept[1], signal one occurs after interaction between the antigen and the Ig receptor and results in an irreversible state of immunological tolerance. Only the simultaneous addition of both signals results in lymphocyte activation, whereas signal two alone has no effect. The alternative hypothesis suggests that only one non-specific signal activates the B lymphocytes[5]. This signal is not delivered by the Ig receptors, but rather by non-clonally distributed receptors for PBA, since B lympho-

Fig. 1 Induction of DNA synthesis with different concentrations of LPS in bovine lymphocytes in serum-free microcultures using 5×10^5 cells per well in 0.2 ml of Eagle's minimum essential medium in Earle's solution containing amino acids and antibiotics as described by Mishell and Dutton[9,11]. The cultures were incubated at 37 °C as described before[12], and after 1–2 d each well received 1 µCi of ^{3}H-thymidine. The cultures were then collected 24 h later on glass fibre filters in a multiple harvester. The filters were dried, scintillation fluid was added and they were counted in a liquid scintillation counter. The figures below the bars denote the cow number in Table 1.
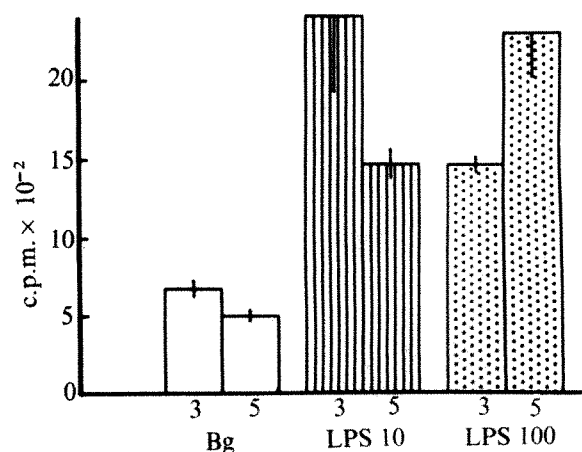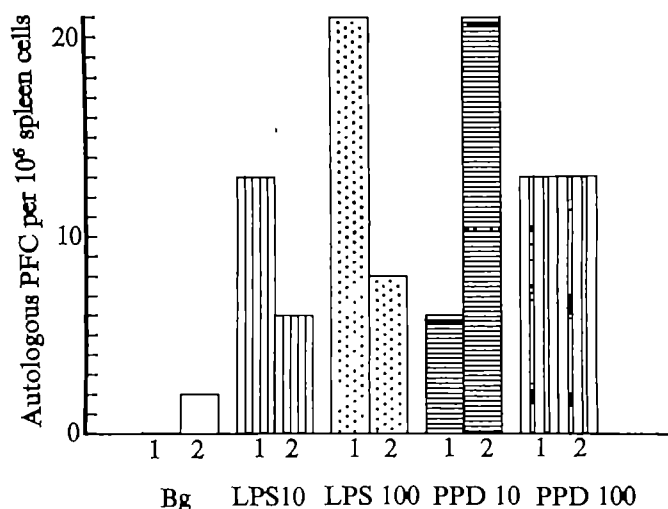
Fig. 2 Induction of autologous PFC in serum-free cultures with bovine lymphocytes stimulated with different concentrations of LPS and PPD. The numbers below the bars indicate cow number. Values are expressed as PFC per 10⁶ lymphocytes. 1×10^7–2×10^7 spleen cells were cultured in 1 ml of medium in 3-cm plastic Petri dishes set up in triplicate The cultures were incubated on a rocking platform for 2 d and thereafter the cells were assayed in a modified[13] agar plaque assay[14] using untreated autologous red cells or haptenated (NNP) sheep red cells. Haptenation was carried out as before[13].

cytes from bovine spleens must necessarily have been in continuous contact with their own red cells, and in spite of this did not become irreversibly tolerant. The results suggest that interaction between Ig receptors and self antigens does not give a tolerogenic signal.

It has been pointed out that the self–non-self discrimination might be executed entirely at the T-cell level[2-5]. This is usually sufficient to maintain a state of self–non-self discrimination, because the autologous antigens are thymus dependent and thereby require cooperation between T and B cells for induction of an immune response. Since the self-reacting T cells would be absent, induction would not usually occur. Induction could be achieved, however, if the cells are given the relevant triggering stimulus delivered by different polyclonal B-cell activators as shown in this experimental system. This would usually occur as a consequence of exposure to PBA material, as happens in infections with organisms having or releasing PBA material. This is probably the reason for autoantibody formation in chronic infectious diseases such as mycoplasma infections,

leprosy, malaria and so on, where it has been shown clearly that the infectious organisms have polyclonal B-cell activating properties.

LENNART HAMMARSTRÖM
EDVARD SMITH
DANIELE PRIMI
GÖRAN MÖLLER

Division of Immunobiology,
Karolinska Institutet,
Wallenberglaboratory, Lilla Frescati,
104 05 Stockholm 50, Sweden

1 Brotscher, P A , and Cohn, M S., *Science*, 169, 1042 (1970)
2 Weigle, W. O , Chiller, J M , and Habicht, G 8 , *Transplant. Rev.*, 8, 3 (1972).
3 Bankhurst, A D., Torrigiani, G , and Allison, A. C., *Lancet*, 1, 226 (1973).
4 Coutinho, A., and Moller, G , *Scand J. Immun* , 4, 99 (1975)
5 Coutinho, A , and Möller, G , *Scand. J Immun* , 3, 133 (1974)
6 Weigle, W D., and Nakamura, R. M., *J Immun.*, 99, 223 (1967).
7 Moller, G., *Nature*, 207, 1166 (1965)
8 Coutinho, A., Gronowicz, E., Bullock, W., and Moller, G , *J. exp. Med.*, 139, 74 (1974).
9 Gronowicz, E., and Coutinho, A., *Eur J. Immun* , 5, 413 (1975).
10 Moller, G., *et al.*, *J. exp. Med* 143, 1429 (1976).
11 Mishell, R L., and Dutton, R W., *J exp Med.*, 126, 423 (1967)
12 Coutinho, A , and Moller, G., *Eur J Immun* , 3, 531 (1973).
13 Bullock, W , and Moller, B , *Eur. J. Immun* , 2, 514 (1972).
14 Jerne, N. K., and Nordin, A A , *Science*, 140, 405 (1963).

Genetic control of the response of chicken leukocytes to a T-cell mitogen

THE way in which mitogens stimulate lymphocytes to divide in culture is not known. Some mitogens (for example, bacterial lipopolysaccharides) stimulate immunoglobulin-bearing lymphocytes to divide and secrete antibodies; others, such as concavanalin A (con A) or phytohaemagglutinin (PHA), stimulate thymus-derived cells (T cells) to proliferate[1,2]. An analysis of the mechanism of action of these mitogens may be helpful in deciding how specific lymphocytes respond to

Table 1 Induction of PFC against NNP-coupled and autologous red blood cells by LPS after 4 d in serum-free cultures containing 2×10^7 cells each

Cow no.	Target cells	0	PFC per culture after addition of LPS		
			10 µg	100 µg	1,000 µg
3	NNP-SRC	32	17	167	
	Autologous	0	0	7	
4	NNP-SRC	3	5	34	
	Autologous	0	1	6	
5	NNP-SRC	63	121	60	
	Autologous	0	5	3	
6	NNP-SRC	46	60	108	80
	Autologous	2	4	14	8
7	NNP-SRC	7	32	27	
	Autologous	6	24	31	

Bovine spleens were collected immediately after slaughter and the interior part was cut out. Blood samples were drawn simultaneously from the same animal.

Fig. 1 Responses of peripheral blood leukocytes from CB (·--·), WA (——), and (CB × WA)F₁ (--) birds to various doses of con A. Mean and range of responses of four birds from each strain are given for each dose The response of a single bird is measured as the mean of ³H-thymidine uptake in triplicate 96-h microcultures, each culture containing 1.5×10^6 leukocytes in 200 µl RPMI 1640 medium supplemented with penicillin (100 IU ml⁻¹), streptomycin (100 µg ml⁻¹), and glutamine (2 mM) plus the specified dose of con A (Miles, Kankakee, Illinois). Cultures were set up in flat-bottomed wells of Falcon 3040 plates and collected as described previously[6] 16 h after addition of ³H-thymidine (2 µCi in 50 µl culture medium; specific activity 2 Ci mmol⁻¹) to each culture.

antigen. We report here a case of genetic control of the ability of leukocytes to respond to a T-cell mitogen; lymphocytes from certain inbred chicken lines give a high proliferative response to con A, whereas those of other lines give a low response. The chicken provides a suitable model for studying such responses, in particular because it is easy to culture avian peripheral blood leukocytes and thus compare the results of mitogen stimulation directly with those obtained using human peripheral blood cells.

The CB and WA chicken lines used in these experiments are highly inbred[3,4]. Peripheral blood leukocytes, prepared as described by Greaves et al.[5], were cultured in the presence of varying doses of con A and PHA in microculture trays; after a fixed period (usually 3 d) ^{3}H-thymidine was added to the cultures and after a further 16 h, the cells were collected and uptake of radioactivity (presumed to reflect DNA synthesis) was determined as described previously.[6].

Figure 1 shows the responses of leukocytes from adult (9–36-month-old) birds of the CB and WA strains, and of $(CB \times WA)F_1$ hybrid birds, to various doses of con A. The responses of WA and F_1 birds are similar and are at least fivefold higher than the response of CB birds at all doses of con A except the (possibly toxic) highest tested. The responses were measured after 96 h, but, as Fig. 2 shows (for a selected dose of con A), the difference between the two strains can be seen throughout the response (that is, from 2–4 d).

Figure 2 also shows that leukocytes of the two strains respond equally well, or almost so, to a second T-cell mitogen, PHA; although only shown for a particular dose of PHA, this is true for a wide range of PHA concentrations. This fact, as well as providing a useful control for the survival of the leukocytes in culture, shows that the low response of the CB strain to con A is not caused by a defect in the ability of CB leukocytes to proliferate in culture (indeed, as shown previously, the response of CB leukocytes in a mixed leukocyte reaction is as good, or almost as good, as that of WA leukocytes[6]).

An analysis of the responses to con A of leukocytes from $(CB \times WA)F_2$ and $CB \times (CB \times WA)$ backcross birds is given in Fig. 3. Of 31 F_2 birds tested, 22 were high and 9 low responders. Of 8 backcross birds, 5 were high and 3 low responders. Most, including all the low responders, have been tested at least twice

Fig. 3 Genetic control of the response of chicken peripheral blood leukocytes to con A. The responses of leukocytes from individual CB ($\bigcirc$), WA ($\bullet$) and (CB × WA)F$_2$ ($\blacksquare$) or CB × (CB × WA)F$_1$ backcross ($\blacksquare$) birds were measured after 96 h in culture, as described in the legend to Fig. 1, in three experiments. In the first (a), the responses of CB, WA and 14 randomly chosen (CB × WA)F$_2$ hens were measured; of the F$_2$ birds, 3 are low, and 11 high responders. In the second experiment (b), the responses of CB, WA and 14 selected F$_2$ birds (hens $\female$ and cocks $\male$) were measured; of the F$_2$ birds, 7 are high and 7 low responders. The birds were selected as high and low responders on the basis of previous experiments. In the third experiment (c), the responses of CB, WA, and 8 randomly chosen CB × (CB × WA) backcross birds were measured; of the backcross birds, 5 are high and 3 low responders. In each experiment, the responses to various doses of con A were measured; the values given are responses to the dose giving the optimum distinction between high and low responders (2.5, 2.5 and 0.25 μg ml^{-1} of con A, respectively).

with consistent results. These findings indicate that a single, dominant gene controls the ability of peripheral blood leukocytes from the two strains to respond to con A in culture. We have provisionally called this locus *Mr*1 (mitogen response 1) and its alleles *Mr*1-*hi* (WA) and *Mr*1-*lo* (CB).

Table 1 gives the results of tests for genetic linkage of the *Mr*1 locus to the major histocompatibility (B) complex (which actually includes at least three closely linked loci[7–9]), and to the linked G1 and M1 immunoglobulin allotype loci[10,11].

Fig. 2 Time course of responses of peripheral blood leukocytes from CB and WA birds to con A (2.5 μg ml^{-1}) and PHA (1:100 diluted from PHA-M (Difco) stock). Cultures were set up in triplicate for each time point and pulsed for 16 hr with ^{3}H-thymidine before being collected at the specified times, as described in the legend to Fig. 1. The source of CB (or WA) cells was a pool of leukocytes prepared from three CB (or WA) birds. Each point represents the mean of the triplicate values. Vertical bars represent the range of responses of three individual CB ([) and WA (]) birds to PHA (1:100 dilution), measured after 96 h (in a separate experiment, in identical conditions).

Table 1 An autosomal locus, not part of the major histocompatibility or immunoglobulin gene complexes, controls the response of chicken peripheral blood leukocytes to con A

	High	Low	Total
	21 (22.5)	9 (7.5)	30
Bl/1	8 (9.0)	4 (3.0)	12
Bl/9	9 (9.7)	4 (3.3)	13
B9/9	4 (3.7)	1 (1.3)	5
Gla/a	5 (4.5)	1 (1.5)	6
Gla/d	8 (10.5)	6 (3.5)	14
Gld/d	8 (7.5)	2 (2.5)	10
$\male$	10 (12)	6 (4)	16
$\female$	11 (10.5)	3 (3.5)	14

Phenotypes of 30 randomly chosen chickens derived from an F$_2$ cross between the inbred lines CB and WA. The B locus antigens (Bl or 9)[3,4], immunoglobulin G allotypes (Gla or d, which in this cross are associated with the immunoglobulin M allotypes Mla or Mlc respectively)[10,11] and sex of each bird were determined, as well as the reactivity (high or low) of its peripheral blood leukocytes towards con A. Numbers of birds of each phenotype are given, together with (in parentheses) expected numbers if there is no genetic linkage between loci controlling response to con A and the other listed factors. Phenotypes of the parental lines are: CB, Bl/1, Gla/a; WA, B9/9, Gld/d.

The *Mr*1 locus is not linked to these loci, nor is it sex linked or sex limited (Fig. 2). Tests for linkage of the *Mr*1 locus to loci controlling a serum albumin polymorphism[12], feather colour[13], and the C and E blood group antigens[14] have also proved negative (J.R.L.P., V.M., and K. Hála, unpublished).

A single-gene effect on the ability of leukocytes to respond to con A could have a number of explanations, of which some may be tested directly. For example, the WA, but not the CB, strain may have a lymphocyte surface receptor which has a high affinity for con A, and which, on reaction with con A, triggers the cell to divide. An alternative, but equally likely, explanation is that a subpopulation of leukocytes (presumably of T lymphocytes[2]), whose size differs in the two strains, responds to con A. The *Mr*1 marker may prove useful in identifying the putative receptor, or T-cell subpopulation, involved in the response.

Other cases of genetic difference in the ability to respond to mitogens have been described: lymphocytes from mice of the C3H/HeJ strain do not respond to lipopolysaccharides, whereas cells from other mouse strains respond well; the response is controlled by a single non-H–2-linked and non-Ig allotype-linked gene[15]. Small differences, shown to be genetic, in the ability of different mouse and rat strains to respond to the T-cell mitogens con A and PHA have been described[16,17]; this differs from our results in that low responders to con A were also low responders to PHA, suggesting that polymorphism of a specific con A receptor is not an explanation for the observed differences. No genetic differences in the ability to respond to mitogens have been established in human populations, but that such differences may exist is clearly a point to consider when the results of mitogen stimulation assays of human T-cell function are being assessed.

V. Miggiano
M. North
A. Buder
J. R. L. Pink

Basel Institute for Immunology,
487 Grenzacherstrasse, Postfach 4005,
Basel 5, Switzerland

1 Moller, G. (ed.), *Transplant. Rev.*, 11 (1972).
2 Weber, W. T., *Transplant. Rev.*, 24, 113 (1975).
3 Hašek, M., Knizetová, F., and Mervartová, H., *Folia biol.*, 12, 335 (1966).
4 Hála, K., Hašek, M., Hlozánek, I., Hort, J., Knizetová, F., and Mérvartová, H., *Folia biol.*, 12, 407 (1966).
5 Greaves, M. F., Roitt, I. M., and Rose, M. E., *Nature*, 220, 293 (1968).
6 Miggiano, V., Birgen, I., and Pink, J. R. L., *Eur. J. Immunol.*, 4, 397 (1974).
7 Schierman, L. W., and McBride, R. A., *Transplantation*, 8, 515 (1969).
8 Hála, K., Vilhelmová, M., and Hartmanová, J., *Immunogenetics*, 3, 97 (1976).
9 Ziegler, A., and Pink, J. R. L., *J. biol. Chem.* (in the press).
10 Ivanyi, J., *Immunogenetics*, 2, 69 (1975).
11 Pink, J. R. L., and Ivanyi, J., *Eur. J. Immun.*, 5, 506 (1975).
12 McIndoe, W. M., *Nature*, 195, 353 (1962).
13 Hutt, F. B., in *Pigment Cell Growth* (edit. by Gordon, M.), 29 (Academic, London, 1953).
14 Gilmour, D. G., in *Physiology and Biochemistry of the Domestic Fowl*, 2 (edit. by Bell, D. J., and Freeman, B. M.), 883 (Academic, London, 1971).
15 Watson, J., and Riblet, R., *J. exp. Med.*, 140, 1147 (1974).
16 William, M., and Benacerraf, B., *J. exp. Med.*, 135, 1279 (1972).
17 Newlin, C. M., and Gasser, D. L., *J. Immun.*, 110, 622 (1973).

Presence of HLA-D
determinants on human macrophages

When lymphocytes from two genetically dissimilar individuals are cultured *in vitro*, blast transformation and increased DNA synthesis occur in some of them; this is the mixed lymphocyte culture (MLC) interaction[1,2]. In man, the MLC-activating determinants seem to be coded for by the genes of the HLA chromosome region, of which the HLA-D determinants seem to induce the strongest proliferation. HLA-D determinants are expressed on B cells[3], skin cells[4] and sperm[5]. Human macrophages can also activate allogeneic lymphocytes *in vitro* as detected in the

mixed lymphocyte macrophage culture (MLMC) interaction[6,7]. The stimulating activity of the macrophages seems to be much greater than that of peripheral blood lymphocytes. Human Ia-like antigens (B-cell antigens) are strongly expressed on macrophages[8,9]. In the mouse, the lymphocyte-activating determinants of the *M* locus as well as those of the *H–2* loci are also expressed on macrophages[10]. We now present evidence that the lymphocyte-activating determinants on human macrophages are also coded for by the *HLA-D* locus and that an antiserum against HLA-Dw2 (Ia-equivalent) or closely associated determinants can inhibit specifically the stimulatory properties of macrophages from Dw2-positive donors.

Leukocytes were separated from human peripheral blood by Ficoll–Isopaque flotation. Relatively pure macrophage populations were obtained by incubating the isolated leukocytes in plastic flasks for 8 h. The non-adherent cells, mainly lymphocytes, were decanted by repeated washing and the adherent cells were incubated for a further 7–8 d, after which the peripheral blood monocytes had matured into macrophages and the granulocytes had died out. The macrophages were collected by scraping the flask sides with a rubber policeman in calcium-free medium containing EDTA, and the cells were washed out in the same medium. The MLMC technique is described in the legend to Fig. 1. Stimulating macrophages were obtained from individuals known to be homozygous for a particular HLA-D determinant and used as typing cells against a panel of lymphocytes previously typed to be HLA-D heterozygous. In addition, macrophages were isolated from HLA-D heterozygotes and used as control stimulators for determination of the relative response. The anti-HLA-Dw2-associated antibody serum SφW was produced by planned immunisation of HLA-A, HLA-B identical volunteers[11].

The pooled results from five experiments with macrophages from donors homozygous for either the Dw2 or Dw3 determinants are shown in Fig. 1. Macrophages homozygous for a particular HLA-D determinant stimulated rela-

Fig. 1 HLA-D typing results obtained with macrophages from HLA-D homozygous donors (Dw2 and Dw3). Responding lymphocytes were separated from donors previously typed to be Dw2 positive (○) or Dw2 negative (■); Dw3 positive (▲) or Dw3 negative (●). Responding cells (50×10^3) were mixed in combination with macrophages (10×10^3) in 0.2 ml of medium RPMI 1640 containing antibiotics and 20% human AB serum in the wells of microtitre plates. The plates were incubated for 5 d and collected after an 18-h pulse of ^{3}H-thymidine. The results are expressed in relative response (RR) calculated from the formula:

$$RR = 100 \times \frac{\text{c.p.m. MLMC (homo.macrophages)} - \text{c.p.m. autologous culture}}{\text{c.p.m. MLMC (heteroz.macrophages)} - \text{c.p.m. autologous culture}}$$

The heterozygous macrophage control donors were always Dw2 or Dw3 negative respectively. Pooled data from five experiments.
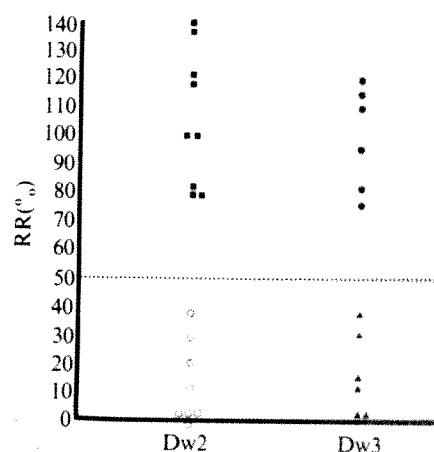
tively weakly (RR < 40%) lymphocytes from heterozygous donors previously typed to possess the HLA-D determinant shared with the macrophage donor. This was the case for both the Dw2 and Dw3 determinants in all combinations tested (Fig. 1). Macrophages did, however, stimulate lymphocytes not sharing the particular HLA-D determinant to a degree comparable with that observed with stimulation by heterozygous macrophages (RR > 70%), where a known two-haplotype incompatibility existed. A clear bimodality of the relative response distributions for the various responding cell donors, similar to that observed using homozygous lymphocytes as stimulating cells in typing MLC tests was observed for the macrophage stimulation for both specificities tested. In no experiment was it possible to demonstrate stimulation of lymphocytes by autologous macrophages, nor to demonstrate increased macrophage proliferation by culturing these cells in combination with mitomycin-treated allogeneic lymphocytes.

It seems therefore that the HLA-D determinants are found in the cell membrane of both lymphocytes and macrophages. This would account for the fact that macrophages from HLA-D homozygous donors failed to, or only feebly stimulated heterozygous lymphocytes sharing an HLA-D determinant. In other words, the HLA-D typing results obtained with homozygous macrophages or lymphocytes were completely congruent, and so a macrophage-specific differentiation antigen cannot be responsible for the positive response observed in MLMCs. In addition, little stimulation was observed in MLMCs between unrelated individuals typed to be HLA-D identical (unpublished observations).

We investigated whether antiserum shown specifically to inhibit the stimulating capacity of lymphocytes possessing the Dw2 determinant could also inhibit specifically the stimulating capacity of macrophages from Dw2-positive donors. This antiserum has been found to inhibit the stimulating capacity of lymphocytes possessing the Dw2 determinant, while inhibiting the responding cells to a much lesser degree[11]. MLMCs were established in medium supplemented with either 20% normal human AB serum or a 1 : 1 mixture of the antiserum in normal AB serum. The antiserum had a significant inhibitory effect (> 50% reduction compared with controls) on all the stimulating macrophages possessing the Dw2 determinant (Fig. 2), but inhibited

Dw2-negative cells only slightly. The inhibiting effect of macrophages from individuals shown to be homozygous for the Dw2 specificity was particularly pronounced (Fig. 2). The stimulating properties of homozygous cells of other specificities however, were not inhibited significantly. The results of these inhibition experiments support the assumption that identical or closely linked stimulating structures (HLA-D) exist on macrophages as well as on lymphocytes.

The macrophage has a well known helper function in MLC, and cultures devoid of such cells do not give a positive response[12]. No specificity for this function has been demonstrated in human MLC, and the necessary macrophages can be either autologous or allogeneic to the responding cell. This helper function can be replaced by cell-free supernatants derived from adherent cells[14], but not by mercaptoethanol[7]. Macrophages are also important in the presentation of mitogens and soluble antigens to T cells[15,16], and in the mouse effective cooperation requires identity between the macrophages and the T cells at the H-2 I region[17], and therefore possibly the Ia determinants. It has been suggested that alteration of Ia determinants by antigen is the stimulatory signal in lymphocyte activation[18]. Our finding that the HLA-D determinants are present on human macrophages may indicate a similar function for HLA-D in man. The ability of macrophages to stimulate allogeneic lymphocytes, however, is probably a result of direct interaction between the stimulating determinants in the macrophage cell membrane and the responding lymphocytes, and not to a helper function *per se*, although this ability may be part of a more general T-cell activation mechanism.

This work was supported by a grant from the Norwegian Research Council for Science and the Humanities.

H. HIRSCHBERG
A. KAAKINEN
E. THORSBY

*Tissue Typing Laboratory,
National Hospital of Norway,
Oslo 1, Norway*

1 Bach, F. H., and Hirschhorn, K., *Science*, 142, 813-814 (1964).
2 Bain, B., Vas, M R , and Lowenstein, L., *Blood*, 23, 108-116 (1964).
3 Lohrmann, H. B., Novikovs, L., and Graw, R L., *Nature*, 250, 144-146 (1974).
4 Hirschberg, H., and Thorsby, E., *Transplantation*, 21, 343-347 (1976).
5 Halim, K , and Festenstein, H , *Lancet*, ii, 1255-1256 (1975).
6 Marshall, W H., Rigo, 8. J , and Melman, 8 , *Lancet*, i, 730-732 (1966).
7 Rode, H. N., and Gordon, J., *Cell. Immun* , 13, 87-94 (1974)
8 Wernet, P , Winchester, R. J , Dupont, B., and Kunkel, H. G , in *Histocompatibility Testing 1975*, 637-642 (Munksgaard, Copenhagen, 1975).
9 Winchester, R J., *et al.*, in *Histocompatibility Testing 1975*, 651-656 (Munksgaard, Copenhagen, 1975).
10 Schirrmacher, V , Pena-Martinez, J , and Festenstein, H., *Nature*, 255, 155-156 (1975).
11 Solheim, B G , Bratlie, A , Winther, N , and Thorsby, E , in *Histocompatibility Testing 1975*, 713-718 (Munksgaard, Copenhagen, 1975).
12 Bondevik, H., Albrechtsen, D., Solheim, B. G , and Thorsby, E , in *Histocompatibility Testing 1975*, 604-607 (Munksgaard, Copenhagen, 1975).
13 Rode, H. N., and Gordon, J., *J Immun* , 104, 1453-1457 (1970).
14 Bach, F H , Alter, B J., Solliday, S , Zoschke, D. C , and Jann, M., *Cell. Immun* , 1, 219-224 (1970).
15 Erb, P , and Feldman, M., *Nature*, 254, 352-354 (1975).
16 Rosenstreich, D. L , Farrar, J. J , and Dougherty, S., *J Immun* , 116, 131-139 (1976).
17 Rosenthal, A S., and Shevach, E. M., *J. exp Med* , 138, 1194-1212 (1973).
18 Sachs, D H., and Dickler, H. B. *Transplant Rev.*, 23, 159-161 (1975).

Fig. 2 The inhibiting effect of an anti-HLA-Dw2-associated antiserum on macrophage stimulation. Open columns, c.p.m. of the cultures in 20% normal serum-supplemented medium; stippled columns, results in medium supplemented with 10% normal serum plus 10% antiserum. *a*, Responder (Dw3/oh) was the antiserum producer; *b*, an unrelated donor (Dw5/Dw6), not possessing the Dw2 determinant. The HLA-D typing data of the macrophage donors are shown. The results are the median and the range of triplicate cultures in c.p.m. Autologous responses have been subtracted.

Origin of immunoglobulin–albumin complexes

THE propensity of human IgA and IgM monoclonal proteins to complex to endogenous serum proteins, such as albumin and α_1 antitrypsin (α_1AT), has been previously established[1-3]. When complexes of IgA-albumin and IgA-α_1 antitrypsin are isolated, albumin and α_1AT are found on separate polymeric IgA molecules and not on monomeric IgA[3]. Another polypeptide chain, J chain, has also been shown to exist in polymeric IgA and IgM (refs 4 and 5) but not monomeric immunoglobulin. Furthermore, J chain is synthesised within IgA and IgM mouse plasma cells[6]. Therefore, it is important to determine

whether albumin is complexed to immunoglobulin as a step in intracellular biosynthesis or whether it is incorporated as a post-secretory event. We report here that IgG and IgA mouse plasmacytoma cells synthesise and secrete albumin complexed to immunoglobulin.

The synthesis of mouse immunoglobulin and albumin was studied using the incorporation of ^{14}C-leucine into plasmacytoma cells of the IgG (LPC-1), IgA (MOPC-315A, TEPC-15, HOPC-8) and IgM (MOPC-104E) class as described pre-

viously[7,8]. All tumour lines were obtained from Litter Biometrics and maintained by serial subcutaneous passage in BALB/c mice. Their identity was confirmed by direct immunofluorescence and by light microscopy. After detergent lysis with 0.5% Nonidet P-40 at 4 °C for 15 min, protein from cell lysates and secretions was extracted using monospecific antiserum to mouse proteins above equivalence. The antisera to mouse IgG and IgA was subjected to affinity chromatography using a mouse albumin immunoadsorbent column. The anti-albumin antiserum was adsorbed using a mouse IgG immunoadsorbent column and shown not to precipitate radioiodinated mouse IgG double antibody precipitation. As a control for nonspecific precipitation, normal rabbit serum and an excess of goat anti-rabbit 7S globulin antiserum was added to an aliquot of all preparations. The immune precipitates were then analysed on sodium dodecyl sulphate (SDS) gels[7].

The results of SDS–acrylamide (5%) electrophoresis of secretions from LPC-1 plasmacytoma cells, collected after 3 h of incubation with ^{14}C-leucine, are shown in Figs 1 and 2. Immunoprecipitation of IgG from these secretions followed by reduction yields two radioactive peaks, corresponding to γ and L chains (Fig. 1a). Complete depletion of IgG from secretions was shown by splitting the supernatant of the IgG precipitate into two fractions and testing aliquots, and reprecipitating with
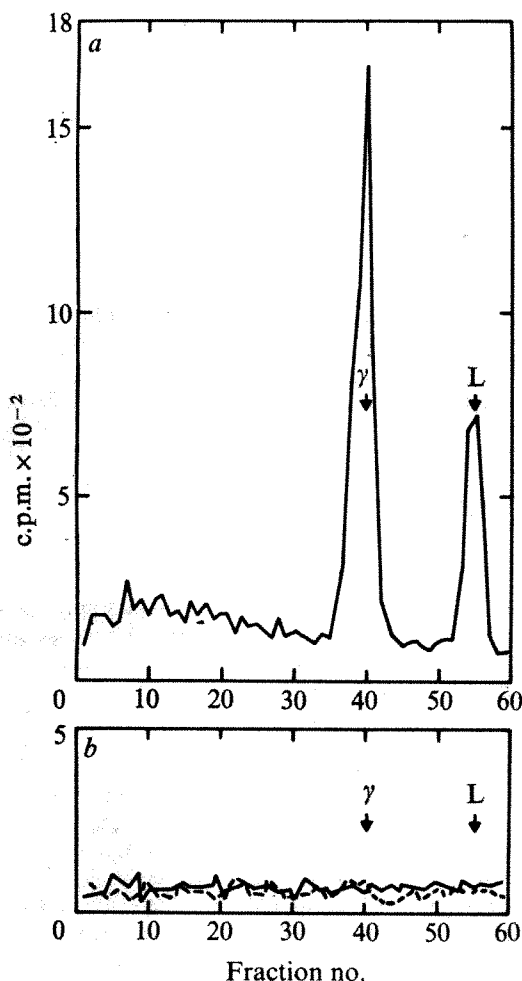
Fig. 1 SDS–polyacrylamide (5%) gel electrophoresis of reduced and alkylated radioactive, secreted mouse IgG (LPC-1). The LPC-1 cell suspension (15×10^6) was incubated in 3.0 ml of Eagle's medium minus leucine containing 10 μCi of ^{14}C-leucine at 325 mCi mmol^{-1} for 3 h at 37 °C. The cell suspension was chilled to 4 °C and made 0.06 M in iodoacetamide. After removal of the cells by centrifugation the supernatant was adjusted to contain 0.5% (w/v) Nonidet P-40 and dialysed against 0.5% Nonidet P-40 in PBS. The resultant secretion was precipitated in 0.06 M iodoacetamide by the addition above equivalence of a monospecific rabbit antiserum to mouse IgG adsorbed with mouse albumin. The precipitate was removed by centrifugation and the supernatant aliquoted into two fractions and further reacted with either rabbit anti-mouse IgG (200 μl) or rabbit anti-mouse albumin antiserum (200 μl) followed by the appropriate amount of mouse IgG (25 μg) or albumin (25 μg) to ensure complete precipitation. The precipitates were washed and dissolved in 2% sodium dodecyl sulphate (0.1 M iodoacetamide), 0.1 M sodium phosphate, pH 7.0, by heating at 100 °C for 2 min. Reduction was carried out with 0.5 M 2 ME by heating for 2 min at 100 °C and alkylation accomplished with excess iodoacetamide. Samples were electrophoresed in 0.1% SDS. The gels were sliced into 1-mm sections, eluted and counted and plotted with the top of the gel on the left-hand side of the figure. a, Reduced anti-IgG immune precipitate; b, supernatant of a reacted with an excess of either anti-IgG (——) or anti-albumin (- - -) antiserum. Arrows indicate marker proteins run simultaneously on separate gels. γ, Gamma heavy chains; L, light chains.

Fig. 2 SDS–polyacrylamide (5%) electrophoresis of ^{14}C-leucine-labelled LPC-1 cell secretions (3 h) precipitated with rabbit anti-albumin antiserum. Incubation with radioactive leucine and preparation of immune precipitate for SDS electrophoresis was as described in Fig. 1. a, Reduced and alkylated anti-albumin precipitate before (——) and after (- - -) adsorption of antiserum with mouse albumin; b, supernatant of the anti-albumin precipitate in a (——) further reacted with either anti-IgG (——) or anti albumin (- - -) antiserum and analysed after reduction and alkylation. Marker proteins are indicated by arrows.
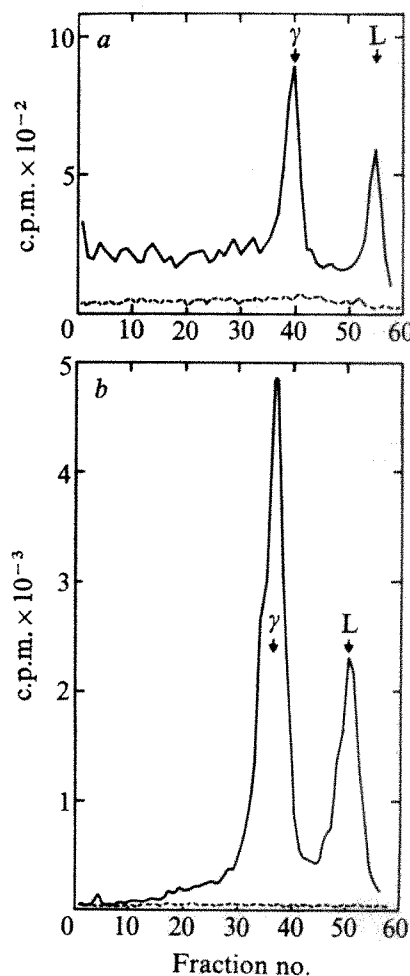
either an anti-IgG or anti-albumin antiserum (Fig. 1*b*). Similar results were obtained using double antibody precipitation with goat anti-rabbit gamma globulin. When anti-albumin antiserum is used to precipitate LPC-1 secretions, two peaks of radio-activity are again seen after reduction, which correspond to γ and L chains (Fig. 2*a*). This radioactive pattern is eliminated when the anti-albumin antiserum is adsorbed with the use of a mouse albumin immunoadsorbent column before precipitation of the secretion (Fig. 2*a*). A control of mouse serum added to secretions and precipitated with rabbit anti-mouse transferrin antiserum showed no precipitated radioactivity on SDS gels. Finally, the supernatant of the anti-albumin precipitate was totally depleted of radiolabelled albumin, while IgG γ and L chains were still present (Fig. 2*b*).

To demonstrate that anti-albumin antiserum specifically precipitated IgG associated with albumin, the amount of IgG (c.p.m.) in secretions before and after this immuno-precipitation was determined. IgG radioactivity in the secretion was depleted by 30–50%, using an anti-IgG antiserum for quantification. These experiments suggest that the precipitation of IgG by anti-albumin antiserum was due to complex forma-tion between albumin and IgG, although no radioactive peak corresponding to albumin was seen on 5% SDS gels.

Since albumin could be secreted unlabelled by the IgG

Fig. 3 LPC-1 cell secretions externally labelled with [125]I. Secretions were collected after 3 h of incubation as in Fig. 1 and then externally labelled[9,10] with ^{125}I and processed as given in legend to Fig. 1. Anti-IgG (*a*) and anti-albumin (*b*) antisera were used to react with secretions and the immune precipitates were electrophoresed, unreduced, on 5% SDS-polyacrylamide gels. Marker proteins, indicated by arrows, were obtained by radio-iodinating isolated mouse IgG and albumin and precipitating these proteins in the presence of unlabelled secretions with monospecific antiserum. Alb, albumin.

Fraction no.
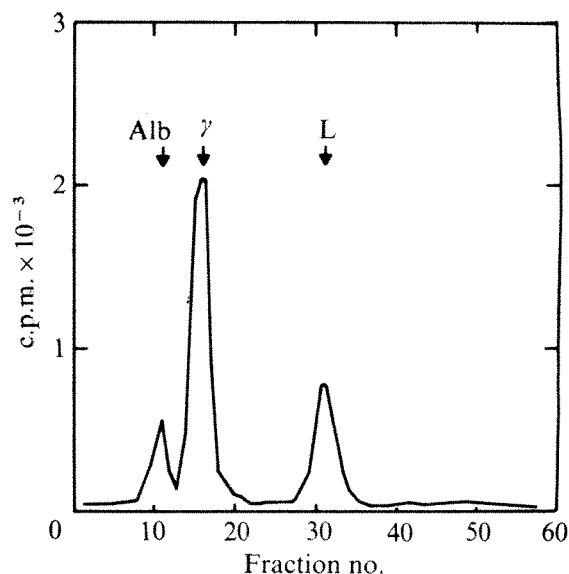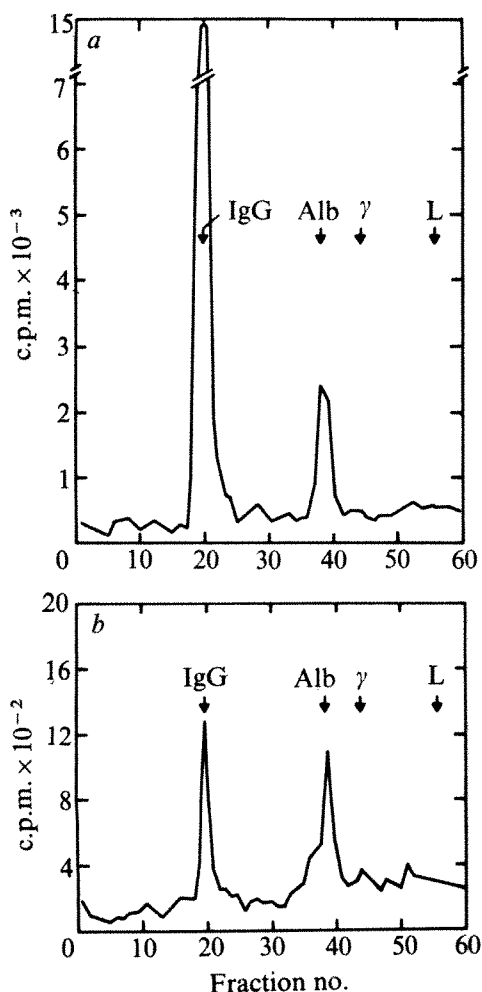
Fig. 4 Radiolabelled albumin secreted from LPC-1 cells after 5 h of incubation with ^{14}C-leucine. Conditions of incubation and handling of immune precipitates were as in Fig. 1. Secretions were precipitated with anti-albumin antiserum and reduced and alkylated and electrophoresed on 10% SDS-acrylamide gels with appropriate markers as in Fig. 1.

plasma cells, LPC-1 cell secretions were externally labelled[9,10] with ^{125}I and after immunoprecipitation were analysed un-reduced on 5% SDS gels. The anti-IgG precipitate of a 3-h secretion had two radioactive peaks corresponding to albumin and IgG (Fig. 3*a*). Similarly, the anti-albumin precipitate showed two peaks with analogous molecular weights on SDS gels (Fig. 3*b*), although the proportion of albumin to IgG differed from that obtained with antiserum to IgG. No discern-ible peaks were obtained on SDS gels of radioiodinated control secretions precipitated with anti-transferrin antiserum or normal rabbit serum and goat anti-rabbit gamma globulin. Planimetry of the graphs showed that the IgG–albumin complex was composed of 1 mol albumin per mol IgG and approximately 30% of the secreted IgG molecules were complexed to albumin. This accounts for the larger IgG peak seen with the IgG pre-cipitate compared to the albumin precipitate. These experiments demonstrate that albumin is non-covalently bound to IgG and does not dissociate in the detergent, Nonidet P-40, although there is dissociation in SDS. Results obtained with three different IgA plasmacytomas were analogous to the IgG plasmacytoma, that is, albumin was secreted non-covalently bound to IgA. Neither albumin synthesis or secretion was found in IgM plasma cells.

To determine whether the albumin secreted from IgG and IgA plasma cells originated after biosynthesis or was adsorbed to the cell, lysates were examined after 3 h incubation with ^{14}C-leucine. In contrast to secretions, the cell lysates demonstrated a distinct radiolabelled peak for albumin, in addition to IgG, with anti-albumin antiserum. Since IgG and albumin were both radiolabelled in the complex present within the cell, it was pos-sible that longer incubation periods would show ^{14}C-leucine incorporated into secreted albumin. The experiment in Fig. 4 depicts a 5-h LPC-1 secretion precipitated with antiserum to albumin and shows a third peak corresponding to albumin, in addition to γ and L chains. Furthermore, gel filtration of 5-h secretions on Biogel P200 equilibrated in 8 M urea demon-strated a component with a molecular weight of 68,000, anti-genically identical to albumin. Similar experiments in three different IgA plasmacytomas demonstrate biosynthesis of albu-min in these cells.

The complexing of albumin to IgG deserves further comment. Since no additional IgG or albumin is found in the supernatant after IgG is precipitated from secretions, this suggests that all of

the secreted albumin is complexed to IgG. One speculation is that the transport of albumin across certain membranes requires immunoglobulin. Furthermore, the albumin found in the complex is secreted by seniority—the older albumin molecules are secreted before the more newly synthesised (labelled) ones. If these complexes are present in the cell membrane, it is possible that the appearance of the complex in the culture media represents cell membrane turnover rather than direct secretion. The biological significance of the complexed albumin may be in the stabilisation of specific receptors on the cell membrane.

STEPHEN P. HAUPTMAN
GLORIA SOBCZAK

Department of Medicine,
Thomas Jefferson University,
Cardeza Foundation,
Philadelphia, Pennsylvania 19107

1 Heremans, J. F., *Les Globulins Seriques du Systeme Gamma* (Editions arcia, Brussels, 1960).
2 Mannik, M., *J. Immun.*, **99**, 899–906 (1967).
3 Tomasi, T. B., Jr, and Hauptman, S. P., *J. Immun.*, **112**, 2274–2277 (1974).
4 Halpern, M. S., and Koshland, M. E., *Nature*, **228**, 1276–1278 (1970).
5 Mastecky, J., Zikan, J., and Butler, W. T., *Science*, **171**, 1163–1165 (1971).
6 Parkhouse, R. M. E., *Nature new Biol.*, **236**, 9–11 (1972).
7 Laskov, R., and Scharff, M. D., *J. exp. Med.*, **131**, 515–541 (1970).
8 Laskov, R., Lanzerotti, R., and Scharff, M. D., *J. molec. Biol.*, **56**, 327–339 (1971).
9 Greenwood, F. C., and Hunter, W. M., *Biochem. J.*, **89**, 114–123 (1963).
10 McConahey, P. J., and Dixon, F. J., *Int. Archs Allerg. appl. Immun.*, **29**, 185–189 (1966).

Solitary cells and enzyme exchange in tetraparental mice

I REPORT here studies on tetraparental mice made with a greatly improved histochemical method for the enzyme β-glucuronidase. There were two unexpected findings: first, small clusters and solitary cells were found in many mosaic tissues of the tetraparental mice, and second, there was evidence that enzyme exchange occurs normally among cells.

Tetraparental mice are made by aggregating pairs of eight-cell embryos *in vitro*[1,2], and those I used were made from pairs of inbred strains whose tissues are known to differ in β-glucuronidase activity[3]. Thus the mosaic tissues of the tetraparental mice were composed of two intermingled populations of cells which could be distinguished microscopically in tissue sections stained histochemically for β-glucuronidase as originally demonstrated by Condamine *et al.*[4].

Tetraparental mice (provided by Richard J. Mullen) were prepared from C57BL/6J or BALB/cWt (high glucuronidase) and C3H/HeJ or AKR/J (low glucuronidase) embryos[5]. Postnatal tetraparental mice 4 d, 9 d, 25 d, 10 months and 16 months old were used. Tissues from these mice and from homozygous controls were treated in parallel. Tissue specimens fixed in 4% formaldehyde were prepared for microscopy by a method involving several unusual steps, including incomplete dehydration, embedding in polyethylene glycol 400 distearate wax treated with NH₄HCO₃ and water, and staining of sections in a solution containing gelatin as well as the usual substrates (details in Fig. 1). This procedure permitted staining and demonstration of mosaicism in a much wider range of cell types than is possible with the method of Condamine *et al.*[4]. In many cells the staining was observed throughout the cytoplasm as well as in the lysosomes. This is in accord with findings that β-glucuronidase, which is a lysosomal enzyme, is also present outside the lysosomes[6,7].

In many tissues of tetraparental mice two populations of cells were readily distinguished by differences in staining intensity. The more intensely stained cells, which I shall call G cells, are assumed to be homozygous for *Gus*ᵇ, the wild-type allele for β-glucuronidase, and thus are assumed to arise from the C57BL/6 or BALB/c embryo. The more lightly stained cells, termed g cells, are assumed to be homozygous for *Gus*ʰ, the allele resulting in lower glucuronidase activity[3] and lower thermal stability[8] of the enzyme; these cells are assumed to arise from the C3H or AKR embryo. Those tissues in which cells were not readily separable into two groups on the basis of staining intensity will be discussed later in this report.

The arrangement of G cells and g cells in a wide range of tissues revealed a thorough intermixture of both types, and a fairly widespread occurrence of small clusters of cells in many tissues (Figs 1–3). (I use 'cluster' to refer to a group of cells of one staining type surrounded by cells of the contrasting type.) There was striking variation in the ratio of the number of G cells to g cells as well as in size of clusters from animal to animal, from organ to organ within the same animal, and even from place to place within the same organ. Though large clusters were common, cells in small clusters (those containing fewer than 10 cells as viewed in a single section) made up between roughly 1% and more
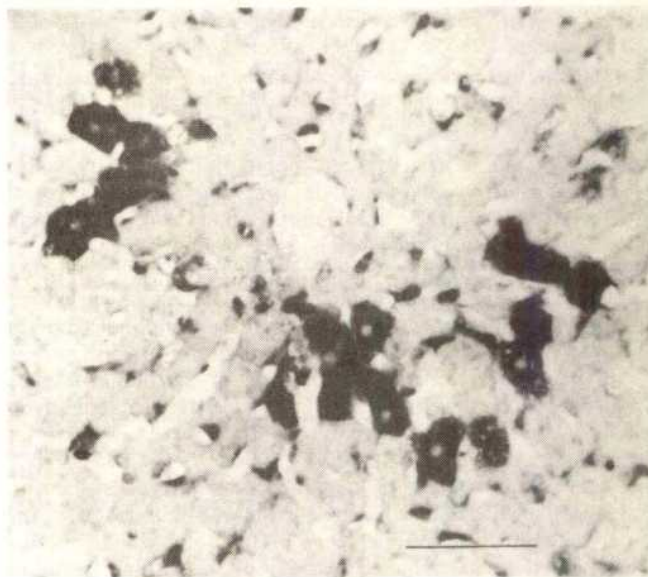
Fig. 1 Liver of 10-month-old tetraparental mouse showing mosaicism after staining for β-glucuronidase. About a dozen hepatocytes are intensely stained (G cells); the remaining hepatocytes—about 90% of all hepatocytes in this field—are lightly stained (g cells). (The much smaller, intensely stained cells are Kupffer cells.) Mice were fixed by intracardiac perfusion for 5 min with 4% formaldehyde (from paraformaldehyde) in a 0.15 M sodium phosphate buffer, pH 7.4. The whole mouse was then placed in a water-ice slurry. After 2–3 h, the organs were removed in a cold room at 4 °C and immersed in the same fixative at 0 °C for a total fixation time of 6 h. Specimens were rinsed at 0–4 °C for 2–5 d in several changes of 6% polyvinyl pyrrolidone (molecular weight 10,000) plus 10% sucrose (compare ref. 24), then dehydrated for 1–2 d at 0–4 °C in 98:2 acetone–water (several changes), and finally embedded in wax. Polyester wax (= polyethylene glycol 400 distearate)²⁵ at 40–50 °C was stirred for 3 h with 5% by weight of solid NH₄HCO₃ and filtered. Two millilitres of water at 80–90 °C were added with stirring to 100 ml of melted wax at 40–50 °C. Tissue specimens were embedded in the treated wax after infiltration for 0.5–2 d at 37–40 °C (several changes). Sections cut at 5–20 μm were mounted in a shallow pool of water on slides treated specially²⁶ with Mayer's albumen. After sections had dried for ~1 d, slides were dewaxed in acetone, then allowed to dry twice from 9:1 acetone–water (to promote section adhesion). Slides were stained at 37 °C in a solution of hexazotised pararosanilin, naphthol AS–BI glucuronide, and 5% gelatin (Eastman No. 1099); with gelatin, periods of staining up to 5 h were possible before a change of staining solution was necessary. Apart from the gelatin, the composition of the solution was the same as that described by Hayashi *et al.*²⁷. Because of the enormous range of glucuronidase concentrations, optimum staining time ranged from 20 min for some cell types (for example, tissue macrophages)²⁶ to 25 h for others (for example, retinal neurones). The bar represents 50 μm.
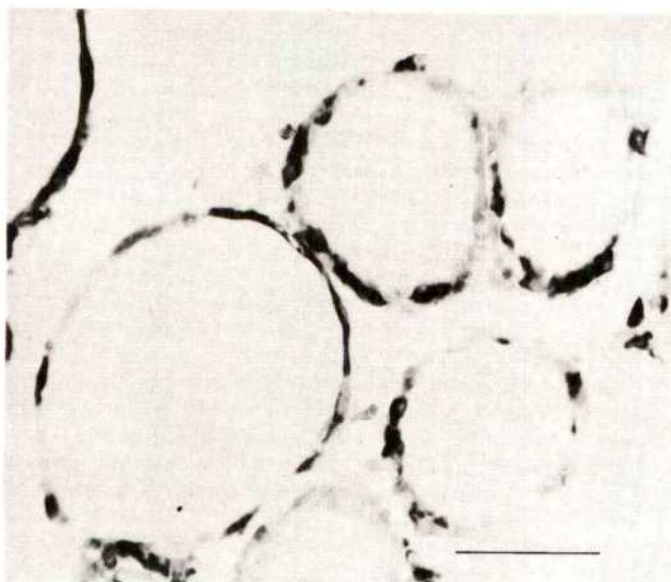
Fig. 2 Thyroid of 16-month-old tetraparental mouse. There are about a dozen G cells and about a dozen g cells in each of the four main follicles. The bar represents 50 µm.

than 20% of the total cell population in many tissues. Small clusters were present in the following tissues of the 16-month-old tetraparental mouse and also in the same tissues of at least one of the three young mice (4, 9, and 25 d old): in the parenchyma of liver, pancreatic islets, exocrine pancreas, mucous and serous part of salivary gland, thyroid, parathyroid, adrenal cortex, seminal vesicle and Harderian gland; in the surface epithelium of salivary ducts, non-glandular stomach (basal layer of epithelium), glandular stomach, pylorus, small intestine, trachea, epididymis, choroid plexus, and ependyma; and in some nerve cell groups in the central nervous system.

The smallest 'clusters' consisted of single cells. The proportion of these cells ranged from a small fraction of a per cent up to a few per cent of the total cell population. Individual cells which seemed to be isolated in two dimensions when viewed in a single section were observed in most tissues listed above. The presence of truly solitary cells (those isolated in three dimensions) was verified by examination of serial sections of liver, salivary gland and duct, small intestine, thyroid, seminal vesicle, epididymis, Harderian gland and choroid plexus. Although small clusters and solitary cells have been noted previously[8], this finding was unexpected: there have been several reports of much larger clusters in mosaic tissues—in the range of 10^4 to 10^6 or more cells per cluster[9-14]. This earlier work was based on methods of tissue analysis too crude for the detection of the small clusters which undoubtedly were present.

The presence of a solitary cell in a mosaic tissue means either that a pair of sister cells have separated after the cell division which produced them or that one member of the pair has died. The solitary cells found in this study may have migrated away from their sister cells at any stage of foetal or postnatal life; my observations provide no evidence on when migration occurred. One can, however, infer that cell migration in the region of large clusters was very slow[15] —no more than a few cell diameters of total migration by the average cell throughout postnatal life.

In the tissues discussed above the G cells and g cells formed two distinct populations readily distinguished by their staining intensities. Unfortunately, as previously mentioned, in some tissues many of the cells were of intermediate staining intensity and could not be assigned to either the G cell or g cell group. These included most of the lymphatic system, most connective tissues and many nerve cell groups in the central nervous system.

The differences in staining intensity between cells of tetraparental mice and cells of the same tissues in homozygous control mice showed some interesting regularities. The G cells of some tissues were stained less intensely in tetraparental mice than in the high-glucuronidase strains (C57BL/6 and BALB/c); this was observed in mucous and duct cells of salivary gland, stomach glandular epithelium, absorptive epithelium of small intestine, colon glandular epithelium, thyroid epithelium, adrenal cortical epithelium, renal tubular epithelium, Harderian gland epithelium, chondrocytes, smooth muscle cells, brown fat cells, some nerve cell bodies in the central nervous system, and Schwann cells (in trigeminal ganglion). Also the g cells of some tissues were stained more intensely in tetraparental mice than in the low-glucuronidase strains (C3H and AKR); this was observed in stomach glandular epithelium, absorptive epithelium of small intestine, colon glandular epithelium, adrenal cortical epithelium, renal tubular epithelium, epididymal epithelium, brown fat cells and choroid plexus.

These observations can be explained most simply by postulating a partial exchange of enzyme between G cells and g cells in tetraparental mice. An alternative explanation is that both the G cells and g cells of tetraparental mice have been influenced by the presence of cells of the opposite type to alter their level of endogenous enzyme. The hypothesis of enzyme exchange is supported by the preliminary finding that some g cells of tetraparental mice contain both Gus^b and Gus^h enzyme[16]. Evidence against the exchange of several non-lysosomal enzymes comes from the failure of attempts to detect hybrid molecules of these enzymes in a wide variety of tissues in tetraparental mice. Since hybrids could not be detected (except in skeletal muscle), the monomers of these enzymes apparently do not move from cell to cell (reviewed in ref. 17).

Change of phenotype by transfer of a substance from cell to cell in a mosaic tissue is a well known phenomenon. In the first recognised instance (described by Sturtevant in 1920; reviewed in ref. 18), the substance transferred between cells and found to be responsible for the observed effect proved to be kynurenine, a metabolite of low molecular weight. The findings reported here may be the first demonstration of exchange of an enzyme between cells in an intact animal.

Intercellular exchange of enzymes, particularly of

Fig. 3 Seminal vesicle of 16-month-old tetraparental mouse. There are two layers of secretory epithelium (simple, low columnar) separated by smooth muscle (almost unstained). One epithelial layer contains only g cells; the other layer, G cells and g cells. The bar represents 50 µm.
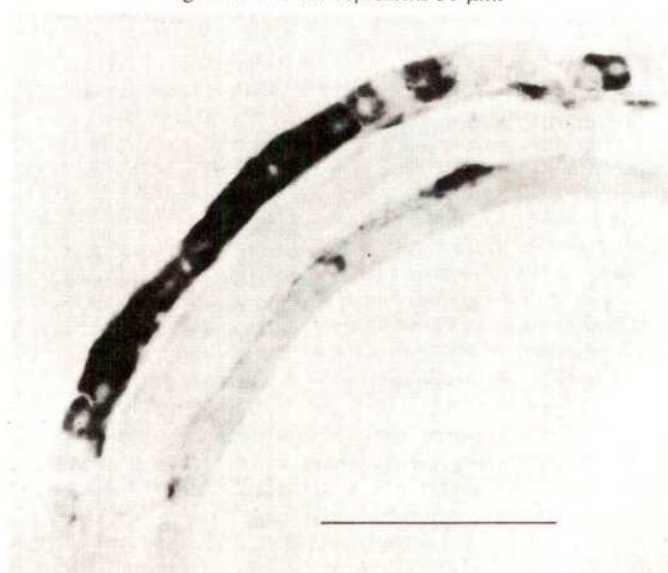
lysosomal enzymes, may occur widely in tissues of ordinary animals. A similar suggestion has been made previously on other grounds[19]. Since many enzymes (including β-glucuronidase) are readily detectable in plasma and other extracellular fluids, these enzymes are available to be taken up by cells. Furthermore, the uptake of exogenous lysosomal enzymes administered experimentally is known to occur both in the intact animal[20] and in cell culture[21]; β-glucuronidase is among the enzymes for which this process has been demonstrated[22,23]. Moreover, the incorporated enzyme has been shown in some instances to assume its normal function inside the cells that take it up[20,21]. These observations, together with the findings presented here, support the view that exchange of glucuronidase (and other enzymes) among cells actually occurs as a normal function in a wide range of cell types.

The idea of using β-glucuronidase as a cell marker in tetraparental mice for studies of embryonic development was suggested to me by Richard L. Sidman.

NED FEDER

National Institute of Arthritis, Metabolism, and
Digestive Diseases,
National Institutes of Health,
Bethesda, Maryland 20014

1 Tarkowski, A. K., *Nature*, 190, 857–860 (1961).
2 Mintz, B., *Am. Zool.*, 2, 432 (1962).
3 Morrow, A. G., Greenspan, E. M., and Carroll, D. M., *J. natn. Cancer Inst.*, 10, 657–661 (1949); 10, 1199–1203 (1950).
4 Condamine, H., Custer, R. P., and Mintz, B., *Proc. natn. Acad. Sci. U.S.A.*, 68, 2032–2036 (1971).
5 Mullen, R. J., and Whitten, W. K., *J. exp. Zool.*, 178, 165–176 (1971).
6 Paigen, K., *Expl Cell Res.*, 25, 286–301 (1961).
7 Fishman, W. H., Ide, H., and Rufo, R., *Histochemie*, 20, 287–299 (1969).
8 Deol, M. S., and Whitten, W. K., *Nature new Biol.*, 238, 159–160 (1972).
9 Ford, C. E., Evans, E. P., and Gardner, R. L., *J. Embryol. exp. Morph.*, 33, 447–457 (1975).
10 Gartler, S. M., Gandini, E., Hutchison, H. T., Campbell, B., and Zechhi, G., *Ann. hum. Genet.*, 35, 1–7 (1971).
11 Linder, D., and Gartler, S. M., *Science*, 150, 67–69 (1965).
12 Linder, D., and Gartler, S. M., *Am. J. hum. Genet.*, 17, 212–220 (1965).
13 Mintz, B., *Proc. natn. Acad. Sci. U.S.A.*, 58, 344–351 (1967).
14 Wegmann, T. G., *Nature*, 225, 462–463 (1970).
15 Lewis, J., *J. theor. Biol.*, 39, 47–54 (1973).
16 Herrup, K., Mullen, R. J., and Feder, N., *Fedn Proc.*, 35, 1371 (1976).
17 Mintz, B., *A. Rev. Genet.*, 8, 411–470 (1974).
18 Ephrussi, B., *Q. Rev. Biol.*, 17, 327–338 (1942).
19 Hickman, S., and Neufeld, E. F., *Biochem. biophys. Res. Commun.*, 49, 992–999 (1972).
20 Desnick, R. J., Krivit, W., and Fiddler, M. B., in *The Prevention of Genetic Disease and Mental Retardation* (edit. by Milunsky, A.), 317–342 (Saunders, Philadelphia, 1975).
21 Neufeld, E. F., in *Cell Communication* (edit. by Cox, R. P.), 217–231 (Wiley, New York, 1974).
22 Hall, C. W., Cantz, M., and Neufeld, E. F., *Archs Biochem. Biophys.*, 155, 32–38 (1973).
23 Lagunoff, D., Nicol, D. M., and Pritzl, P., *Lab. Invest.*, 29, 449–453 (1973).
24 Suzuki, S., *Jap. J. exp. Med.*, 36, 509–517 (1966).
25 Sidman, R. L., Mottla, P. A., and Feder, N., *Stain Technol.*, 36, 279–284 (1961).
26 Feder, N., *J. Histochem. Cytochem.*, 10, 341–347 (1962).
27 Hayashi, M., Nakajima, Y., and Fishman, W. H., *J. Histochem. Cytochem.*, 12, 293–297 (1964).

Hairbulb tyrosinase activity in oculocutaneous albinism

HUMAN oculocutaneous albinism (OCA), characterised by hypopigmentation of skin, hair and eyes, represents a heterogeneous group of at least four distinct autosomal recessive disorders[1,2]. Tyrosinase-negative OCA and tyrosinase-positive OCA are the two most common forms and can be separated by genetic and clinical features[1]. Tyrosinase-negative albinos have no obvious pigment in their skin, hair or eyes, and their plucked hairbulbs form no visible pigment after prolonged incubation in L-tyrosine or L-3,4-dihydroxyphenylalanine (L-dopa). In contrast, tyrosinase-positive albinos have variable, albeit minimal, amounts of pigment in their skin, hair and eyes, and their plucked hairbulbs form large amounts of visible pigment after prolonged incubation in L-tyrosine or L-dopa. The two less common forms of OCA are Hermansky–Pudlak syndrome and yellow mutant OCA. The specific defect in each form of OCA is unknown, and analysis of the biochemical steps in the formation of melanin have been hindered by a

lack of methods suitable for human studies. In this report, we describe an assay for quantifying tyrosinase activity in single human hairbulbs and give the results of the application of this technique to human oculocutaneous albinos.

Melanin formation occurs in the melanocytes, specialised cells found in the skin, hair follicle and eye[3–5]. Tyrosinase catalyses the hydroxylation of tyrosine to dopa and the oxidation of dopa to dopaquinone[6]. Dopaquinone then undergoes non-enzymatic oxidation and polymerisation to form melanin. Our assay measures the tyrosine hydroxylase activity of tyrosinase, determined by the rate of formation of ^{3}HOH with the oxidation of L-3,5-^{3}H-tyrosine to L-5-^{3}H dopa, as described by Pomerantz[7].

Anagen (growing) hairbulbs plucked from the scalp were used as the source of enzyme. Hairbulbs contain a local concentration of melanocytes that remain in the bulb when a hair is plucked[8], and animal studies have shown hairbulb tyrosinase to be present only during the anagen phase of the hair cycle[9,10].

The activity of normal hairbulb tyrosinase was tested from
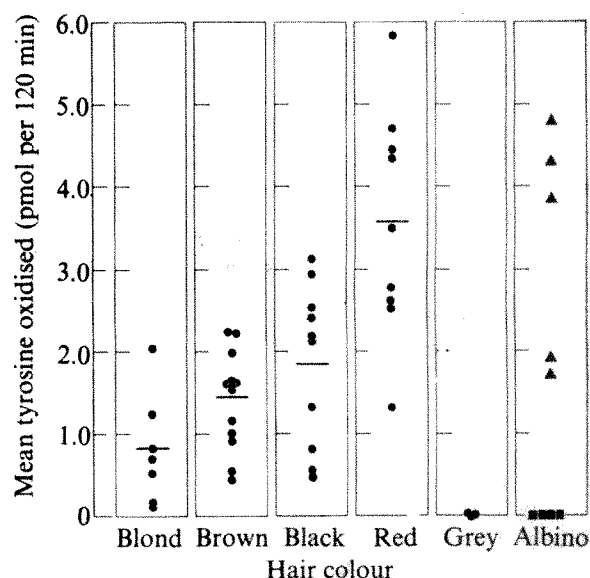
Fig. 1 Mean tyrosine oxidised per 120 min for hairbulbs from normal individuals with different hair colours and from oculocutaneous albinos. For albinos: ■, tyrosinase-negative and ▲, tyrosinase-positive. The overall mean is indicated by a line for blond, brown, black and red hair. Single anagen hairbulbs were identified visually, divided from the shaft at the shaft–sheath junction and incubated in tritiated tyrosine, dopa and Triton X-100. The incubation mixture used was determined to be optimum for single hairbulb tyrosinase quantification after many trials with different buffers, concentrations of substrate and cofactor and incubation volumes. The mixture consisted of L-3,5-^{3}H-tyrosine, 0.2 nmol (1 Ci mmol^{-1}), L-dopa, 0.1 nmol; phosphate buffer, pH 6.8, 6 μmol and Triton X-100, 0.5%, in a total volume of 0.06 ml, and was carried out at 37 °C in a shaking waterbath. Nonspecifically formed ^{3}HOH in the stored L-3,5-^{3}H-tyrosine was eliminated by evaporating the sample of tyrosine to dryness under a stream of nitrogen immediately before use. At 0, 60 and 120 min, a sample of the incubation mixture was placed on a Dowex-50W column equilibrated with 0.1 M citrate. The ^{3}HOH was recovered from the column by washing with 0.8 ml of 0.1 M citrate and counted. The wash removed all tritiated water from the column, with no further activity above background recovered with a second wash. A blank consisting of the incubation mixture with no added hairbulbs was run with each assay. Values for the blank were subtracted from the hairbulb values before determining the amount of tyrosine oxidised. The reaction was linear throughout the 120-min incubation for normally pigmented and tyrosinase-positive albino hairbulbs. Boiled hairbulbs had no activity. Usually six to eight hairbulbs were assayed per subject, with enzyme activity expressed in pmol as the mean tyrosine oxidised per 120 min. The activity in the blanks was equivalent to 0.201 ± 0.152 (mean $\pm$ s.d.) for blond, 0.242 ± 0.089 for brown, 0.312 ± 0.153 for black, 0.255 ± 0.191 for red, 0.163 ± 0.172 for grey and 0.317 ± 0.235 for albino hairbulbs.

forty-two subjects with brown, black, blond, red and grey hair (Fig. 1). All subjects were under the age of 35 except for the three with grey–white hair, and no male was balding. The overall mean for seven subjects with blond hair was 0.821 pmol of tyrosine oxidised per 120 min with a range of 0.134–2.066 pmol per 120 min. Hair colour varied from light white–blond to bright golden blond, but there was no obvious association between degree of 'blondness' and enzyme activity. For thirteen subjects with different shades of medium brown hair a mean of 1.446 pmol of tyrosine was oxidised per 120 min with a range of 0.460–2.267 pmol per 120 min. On repeated testing, hairbulbs from two of the brown-haired subjects oxidised a mean of 1.620 ± 0.312 ($\pm$ s.d.) and 1.660 ± 0.266 pmol of tyrosine per 120 min. Ten subjects with black hair oxidised a mean of 1.881 pmol of tyrosine per 120 min with a range of 0.512–3.162 pmol per 120 min. Two of these subjects were black, and three were Oriental, but there was no association between enzyme activity and racial origin. Hairbulbs from nine subjects with red hair oxidised a mean of 3.596 pmol of tyrosine per 120 min, with a range of 1.352–5.862 pmol per 120 min. The variation was great between these subjects, but there was no obvious association between degree of 'redness' and enzyme activity. Grey–white hairbulbs from three subjects had little or no tyrosinase activity.

Four tyrosinase-negative albinos lacked enzyme activity in their hairbulbs (Fig. 1). Five tyrosinase-positive albinos could be divided according to activity into two groups (Fig. 1). Three had high enzyme activity, oxidising a mean of 4.819, 4.325 and 3.874 pmol of tyrosine per 120 min. These levels are equalled only by normal subjects with red hair. Beard anagen hairbulbs rather than scalp hairbulbs were used from one of the tyrosinase-positive albinos with high activity (4.325 pmol per 120 min) because of the inability to pluck adequate samples of scalp anagen hairbulbs. To verify the use of facial hair, beard hairbulbs from one of the black-haired and two of the brown-haired normal subjects were assayed and were found to have enzyme activity similar to or slightly reduced from the levels found in the same individuals' scalp hairbulbs. Two of the tyrosinase-positive albinos had moderate enzyme activity, oxidising a mean of 1.937 and 1.738 pmol of tyrosine per 120 min. These levels are similar to those for normal subjects with brown and black hair. Clinically, the two with moderate activity differed from those with high activity in that their hair was less pigmented and their visual acuity more reduced. Skin colour was similar in both groups. One of the albinos with moderate activity was 7 months pregnant, but had noticed no changes in skin, hair or eye pigmentation during the pregnancy. Pregnancy affects the hair growth cycle with more hair follicles in anagen[11], but the influence on tyrosinase activity is unknown.

The hydroxylation of tyrosine catalysed by tyrosinase utilises dopa as cofactor, and proceeds very slowly in its absence[6,7]. To show that the hairbulb assay described here is dopa dependent, hairbulbs from a brown-haired subject and from a tyrosinase-positive albino were incubated in different

Table 1 Mean tyrosine oxidised per 120 min for different dopa concentrations in incubation mixture

Dopa concentration (mM)	Mean tyrosine oxidised per 120 min (pmol)	
	Brown	Albino
0	0.000	0.000
1.67×10^{-4}	0.162	0.480
1.67×10^{-3}	1.620	4.325
1.67×10^{-2}	1.883	3.327
1.67×10^{-1}	0.945	—

The albino was a high activity tyrosine-positive albino. The incubation mixture contained 0, 0.01 nmol (1.67×10^{-4} mM), 0.1 nmol (1.67×10^{-3} mM), 1.0 nmol (1.67×10^{-2} mM) or 10 nmol (1.67×10^{-1} mM) of dopa in a total volumn of 0.06 ml. The rest of the assay was as described in the text.

Table 2 Mean tyrosine oxidised per 120 min for hairbulbs from a brown haired subject

Experiment	Mean tyrosine oxidised per 120 min (pmol)
(A) Triton X-100 0.5% (regular assay)	1.620
Triton X-100 2.0%	1.520
Triton X-100 5.0%	0.986
No Triton	0.187
(B) Preincubation in 0.5% Triton	
Hairbulb removed at	
5 min	0.107
10 min	0.721
15 min	1.204
30 min	1.489

For experiment A, hairbulbs were incubated in L-3,5-^{3}H-tyrosine, 0.2 nmol; L-dopa, 0.1 nmol; phosphate buffer, pH 6.8, 2 μmol and Triton X-100, 0–5% in a total volumn of 0.06 ml. For experiment B, hairbulbs were preincubated in 0.03 ml of 0.1 M phosphate buffer, pH 6.8, containing Triton X-100, 0.5%. After the hairbulbs were removed L-3,5-^{3}H-tyrosine 0.2 nmol; and L-dopa, 0.1 nmol were added and the mixture assayed for tyrosinase activity.

concentrations of dopa, as shown in Table 1. When no dopa was added to the incubation mixture there was no oxidation of tyrosine during the 120-min incubation. Dopa concentrations above and below the usual assay concentration gave reduced levels of activity. The assay was clearly dopa dependent. The inhibition at a high dopa concentration was consistent with the action of dopa as a competitive inhibitor of tyrosinase, as shown with mouse and hamster melanoma tyrosinase[6,7].

All the above results were obtained using an incubation mixture containing 0.5% Triton X-100. To test the effect of different detergent concentrations, hairbulbs from a brown-haired subject were assayed in incubation mixtures containing 2%, 5% and no Triton X-100, as shown in Table 2 (experiment A). When no detergent was present, there was little enzyme activity. Triton X-100 (5%) inhibited the activity, and activity with 2% was similar to the regular 0.5%. Triton activation of tyrosinase has been demonstrated, but the exact mechanism of activation is unclear[12]. In the system described here the Triton seemed to release the enzyme from the melanocyte, making it available for assay. This was shown by preincubating brown hairbulbs in 0.5% Triton for 5, 10, 15 and 30 min, after which the hairbulb was removed and the preincubation mixture without hairbulb was assayed for tyrosinase activity by the standard method (Table 2, experiment B). If Triton solubilised the tyrosinase, then the enzyme should have been present in the preincubation solution after the hairbulb had been removed. Hairbulbs preincubated with Triton for 5 min released little enzyme, as shown by an activity of only 0.107 pmol of tyrosine oxidised per 120 min for this preincubation mixture. Preincubation of hairbulbs for 10 and 15 min released progressively more enzyme and preincubation for 30 min released most of the available enzyme as shown by an activity of 1.489 pmol of tyrosine oxidised per 120 min for this pre-incubation mixture. Hairbulbs from this subject oxidised 1.620 pmol, of tyrosine per 120 min in the standard assay.

Three conclusions can be drawn from this study. First, tyrosinase activity can be quantified in single anagen hairbulbs, and varies considerably among normal individuals with different hair colours and among normal individuals with the same hair colour. Although the trend was for more intensely pigmented hair (black and red) to have more hairbulb tyrosinase activity, there was no obvious association between intensity of a specific hair colour and tyrosinase activity. Further studies on the biochemical control of normal hair pigmentation are under way.

Second, tyrosinase-negative albino hairbulbs had no tyrosinase activity by this assay. This is the first quantitative demonstration of absent tyrosinase activity in this form of human OCA and supports the terminology of tyrosinase-negative used by Witkop[1]. The absence of tyrosinase activity

is consistent with the suggestion that this form of OCA is analogous to the c locus (c/c) albino animal[1], for the c locus is thought to control the synthesis of tyrosinase, with c/c animals being deficient in this enzyme[13-15]. Recently, Hearing[12] and Pomerantz and Li[16] have shown that albino mouse eye and skin preparations retain small amounts of tyrosinase activity, 2–3% of that for pigmented eye and skin preparations. Our assay probably would not detect such residual activity in hairbulb preparations. There was definitely no stimulation of activity from tyrosinase-negative albino melanocytes, and there was no evidence for the presence of an inhibitor that could be dissociated from the enzyme by Triton. At present the evidence points to an absence of tyrosinase in tyrosinase-negative OCA.

Third, tyrosinase-positive hairbulbs from tyrosinase-positive albinos contain moderate to large amounts of tyrosinase that can be assayed readily with the system reported here. The defect in this form of OCA does not seem to be an absence of active enzyme, so other mechanisms must be considered. The structure of the tyrosinase-positive melanocyte or melanosome could be altered, but by electron microscopy these appear normal[1]. Substrate availability or transport into the melanocyte could be deficient. The tyrosinase could be structurally abnormal with altered kinetics making it active in non-physiological conditions in vitro but inactive in physiological conditions in vivo. Finally, the activity of tyrosinase could be blocked by an inhibitor, for these have been described for melanoma tyrosinase[17-20]. If an inhibitor is present, the release of tyrosinase from the melanocyte by Triton may also facilitate dissociation of the inhibitor from the enzyme.

There is insufficient information to show which of the above mechanisms is important in the production of tyrosinase-positive OCA. The fact that there seems to be genetic heterogeneity with two types of tyrosinase-positive albinos suggests that more than one mechanism may be active.

I thank J. G. White for review of the manuscript and D. S. Olds for help. This work was supported in part by grants from NIH and the Minnesota Medical Foundation.

RICHARD A. KING

Department of Medicine,
School of Medicine and Division of Human and
    Oral Genetics,
School of Dentistry

CARL J. WITKOP, JR

Division of Human and Oral Genetics,
School of Dentistry,
University of Minnesota,
Minneapolis, Minnesota 55455

1 Witkop, C. J., Jr, Adv. hum. Genet., 2, 61–142 (1971).
2 Witkop, C. J., Jr, White, J. G., and King, R. A., in Heritable Disorders of Amino Acid Metabolism (edit. by Nyhan, W.), 177–261 (Wiley, New York, 1974).
3 Fitzpatrick, T. B., Becker, S. W., Jr, Lerner, A. B., and Montgomery, H., Science, 112, 223–225 (1950).
4 Chase, H. B., Rauch, H., and Smith, V. W., Physiol. Rev., 24, 1–8 (1951).
5 Moyer, F., in The Structure of the Eye (edit. by Smelser, G. K.), 469–486 (Academic, New York, 1961).
6 Lerner, A. B., Fitzpatrick, T. B., Calkins, E., and Summerson, W. H., J. biol. Chem., 178, 185–195 (1949).
7 Pomerantz, S. H., J. biol. Chem., 241, 161–168 (1966).
8 Barnicot, N. A., Birbeck, M. S. C., and Cuckow, F. W., Ann. hum. Genet., 19, 231–249 (1955).
9 Burnett, J. B., Holstein, T. J., and Quevedo, W. C., Jr, J. exp. Zool., 171, 369–376 (1969).
10 Kukita, A., J. invest. Derm., 28, 273–274 (1957).
11 Pecoraro, V., Barman, J. M., and Astore, I., Adv. biol. Skin, 9, 203–210 (1969).
12 Hearing, V. J., Nature new Biol., 245, 81–83 (1973).
13 Foster, M., Adv. Genet., 13, 311–339 (1965).
14 Holstein, T. J., Quevedo, W. C., Jr, and Burnett, J. B., J. exp. Zool., 177, 173–183 (1971).
15 Searle, A. G., Comparative Genetics of Coat Colour in Mammals, 55–115 (Academic, New York, 1968).
16 Pomerantz, S. H., and Li, J. P-C, Nature, 252, 241–243 (1974).
17 Chian, L. T. Y., and Wilgram, G. F., Science, 155, 198–200 (1967).
18 Hamada, T., and Mishima, Y., Br. J. Derm., 86, 385–394 (1972).
19 Romsdahl, M. M., and O'Neill, P. A., in Pigment Cell, 1 (edit. by McGovern, V. J., and Russell, P.), 111–117 (Karger, New York, 1973).
20 Seiji, M., Fukuzawa, H., Mijazaki, N., Akiba, H., and Kato, T., Yale J. biol. Med., 46, 508–515 (1973).

Evidence for inhibin-like activity in bovine follicular fluid

THE negative feedback action of testicular steroids on circulating levels of luteinising hormone (LH) is well known. It is less clear how the testis influences follicle-stimulating hormone (FSH). The existence of 'inhibin', a water-soluble testicular product that inhibits the secretion of FSH, was first postulated in 1932 (ref. 1); recently, inhibin activity has been detected in rete testis fluid[2], testicular tissue[3], spermatozoa[4] and seminal plasma[5]. Immunisation of rabbits with the inhibin fraction of bull seminal plasma has produced an antiserum which raised plasma FSH levels when injected into male and female rats[6]. It has been suggested[7] that, in the male, inhibin may be produced by the Sertoli cells of the testis. If this is the case, one wonders if inhibin could also be produced by granulosa cells of the ovary. We have, therefore, looked for the presence of a specific FSH-suppressing agent in ovarian follicular fluid by comparing the effects of fluid and peripheral plasma on the levels of FSH and LH in newly castrated male rats.

A pool of follicular fluid was collected by aspiration from bovine follicles at various stages of the cycle. The concentrations of testosterone and 17β-oestradiol in the follicular fluid, as measured by radioimmunoassay, were 23.5 and 73.1 ng ml^{-1}, respectively; these values agree with reported levels[8]. Bovine plasma, which has a similar protein composition to follicular fluid[9], was obtained from a castrated bull and testosterone and 17β-oestradiol were added to give concentrations of 25 and 75 ng ml^{-1}, respectively, in a low dose pool of plasma, and concentrations of 250 and 750 ng ml^{-1}, respectively, in a high dose pool.

Five groups of adult male rats (body weight 300–400 g), 6–12 animals per group, were castrated under ether anaesthesia. Approximately 1 ml of blood was collected from the cut spermatic artery at the time of castration (sample 1). Four of the groups of animals were injected intraperitoneally immediately after castration with either 0.25 ml bovine plasma, the low or

Fig. 1 Effect of injection of bovine plasma (BP), BP containing a low dose (LD, follicular fluid concentration) or high dose (HD, ten times follicular fluid concentration) of 17β-oestradiol and testosterone, or bovine follicular fluid (FF) on plasma levels of FSH (a) and LH (b) in male rats 24 h after castration. All values are expressed as the percentage of precastration levels (means ± s.e.m., figures indicate numbers of animals).
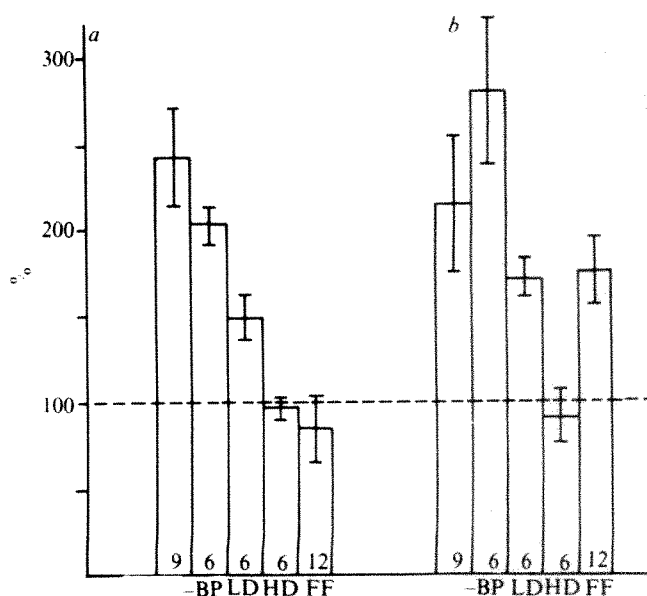
high dose of steroids in plasma, or follicular fluid; and the animals in the fifth group served as untreated controls. Eight hours after castration a second injection of 0.5 ml of the appropriate fluid was administered subcutaneously. The animals were killed by decapitation 16 h after the second injection and blood was collected from the trunk (sample 2). LH and FSH were measured in samples 1 and 2 by radioimmunoassay[10]. The results of these estimations, expressed as the ratio of the concentration in sample 2 to that in sample 1, $\times 100\%$, are shown in Fig. 1.

The untreated controls and the group treated with bovine plasma did not differ significantly; both showed a two- to three-fold increase in plasma LH and FSH levels after castration. Treatment with the low dose of testosterone and 17β-oestradiol caused a significant decrease in plasma LH and FSH compared with bovine plasma-injected groups ($P < 0.05$ and $P < 0.02$, Student's t test); the high dose of steroids caused an even more marked suppression of both gonadotropins to precastration levels. Injection of follicular fluid caused a much greater suppression of plasma FSH than of plasma LH ($P < 0.01$); FSH being reduced to precastration levels, which were significantly lower ($P < 0.05$) than after treatment with the low dose of steroids. In contrast, plasma LH levels were comparable in the follicular fluid and low dose steroid groups. No significant ($P > 0.05$) differences in pituitary LH or FSH levels were detected between any of the five treatment groups.

To establish the non-steroidal, high molecular weight character of the FSH-suppressing agent in the follicular fluid, a second experiment was carried out. Groups of 6–12 animals were injected with steroid-free follicular fluid after castration as described above. LH and FSH were measured in plasma samples collected at the time of castration and 24 h later (Table 1). The effect on FSH levels of injected follicular fluid after ether extraction, charcoal treatment or ultrafiltration (molecular weight > 10,000) was not different from that of the untreated fluid. The suppression of the increase of LH in castrated rats after injection of untreated follicular fluid was abolished when steroids were removed from the fluid by these three treatments; remaining steroid concentrations were < 1% of the original concentrations in all preparations.

These observations suggest that bovine follicular fluid contains inhibin-like activity, which specifically suppresses FSH and cannot be ascribed to steroids, present in the fluid. The active compound has a molecular weight of over 10,000.

Table 1 Effect of injection of steroid-free bovine follicular fluid on plasma levels of LH and FSH in male rats 24 h after castration

Treatment*	n†	FSH	LH
None	12	215±45	302±72
Follicular fluid (FF)	12	99±23‡	160±89‡
FF, ether extracted	6	104±42‡	219±78
FF, charcoal treated	6	99±24‡	294±64
FF (molecular weight > 10,000)	6	90±30‡	274±189
FF (10,000 > molecular weight > 1,000)	6	172±69	247±127

All results are given as the percentage of precastration levels in the same animals (means ±s.d.). Actual levels of FSH and LH at the time of castration were 366±90 (s.d.) ng ml^{-1} NIAMDD rat FSH RP-1 ($n = 48$) and 55.6±22.9 (s.d.) ng ml^{-1} NIAMDD rat LH RP-1 ($n = 48$), respectively.
*FF, Ether extracted: follicular fluid extracted six times with 1.5 volume of ether. FF, Charcoal treated: follicular fluid stirred with charcoal (1 mg ml^{-1}) for 30 min at room temperature and subsequently centrifuged at 12,000g for 30 min. FF (molecular weight > 10,000): follicular fluid diluted to 10 times its original volume with distilled water and subsequently reduced to its original volume by ultrafiltration (Amicon Diaflo PM 10, cutoff point at molecular weight of 10,000). FF (10,000 > molecular weight > 1,000): the volume of the ultrafiltrate of first filtration was reduced to 0.9 times the original volume by ultrafiltration (Amicon Diaflo UM, cutoff point at molecular weight of 1,000).
†Number of animals.
‡Significantly different from untreated controls (Student's t test, $P < 0.05$).

Other indirect evidence, based on the inhibition of follicular development in immature mice injected with bovine follicular fluid[11], leads to a similar conclusion. Since it is unlikely that this substance could be produced by the one oocyte in the follicle[12], the granulosa cells are a more likely source.

It would be attractive to suppose that inhibin might act at the pituitary level, altering the ratio of FSH and LH secreted in response to releasing hormone. This might explain the changing FSH–LH ratio associated with puberty in both the male and female[13] and would give a new dimension to our understanding of the gonadal feedback control of pituitary gonadotropin secretion.

F. H. DE JONG*
R. M. SHARPE

MRC Unit of Reproductive Biology,
2 Forrest Road, Edinburgh, UK

1 McCullagh, D. R., *Science*, 76, 19–20 (1932).
2 Setchell, B. P., and Jacks, F., *J. Endocr.*, 62, 675–676 (1974).
3 Lee, V. W. K., Keogh, E. J., and de Kretser, D. M., *IRCS (Endocr., Physiol., Reprod.)*, 2, 1406 (1974).
4 Lugaro, G., Casellato, M. M., Mazzola, G., Fachini, G., and Carrea, G., *Neuroendocrinology*, 15, 62–68 (1974).
5 Franchimont, P., Chari, S., and Demoulin, A., *J. Reprod. Fert.*, 44, 335–350 (1975).
6 Franchimont, P., Chari, S., Hagelstein, M. T., and Duraiswami, S., *Nature*, 257, 402–404 (1975).
7 Christiansen, P., *Acta endocr. (Kbh)*, 78, 180–191 (1975); De Jong, F. H., and Sharpe, R. M., *Fifth Int. Congr. Endocr.*, Abstract No. 514 (Brühlsche Universitätsdruckerei, Giessen, 1976).
8 Short, R. V., *J. Endocr.*, 23, 401–411 (1962).
9 Caravaglios, R., and Cilotti, R., *J. Endocr.*, 15, 273–278 (1957).
10 Welschen, R., *et al., J. Endocr.*, 64, 37–47 (1975).
11 Peters, H., *Ann. Biol. anim. Biochim. Biophys.*, 13, suppl., 167–176 (1973).
12 Bloom, A. M., and Mukherjee, B., *Expl Cell Res.*, 74, 577–582 (1972).
13 Grumbach, M. M., Grave, G. D., and Mayer, F. E. (eds), *The Control of the Onset of Puberty* (Wiley, New York, 1974).

Prostaglandin mediation of collagenase-induced bone resorption

PATHOLOGICAL bone breakdown may be stimulated by some tissues proximal to bone and the release of collagenase by them may be one cause of this lysis. Inflamed gingivae, for example, have been shown to lyse collagen gels[1,2] and are associated with the loss of bone in periodontal disease. The ability of human mammary carcinomas to break down bone *in vitro*[3,4] has been shown to be caused, in part, by prostaglandin release by the tumours[5] and, in part, by their release of an osteolytically active, non-dialysable material (unpublished results). The possibility that the latter might be collagenolytic in nature was suggested by reports that some neoplasms release collagenolytic activity *in vitro*[6-9]. This prompted our present study of the mechanism of bone breakdown by exogenous collagenase. The investigation has shown that prostaglandin synthesis by bone cells is probably a necessary mediatory step in collagenase-induced bone resorption.

The addition of collagenase to tissue culture medium bathing ^{45}Ca-labelled neonatal mouse calvaria increases their release of ^{45}Ca (Table 1) and results in macroscopically obvious disruption of the calcified matrix, together with connective tissue lysis involving the fronto–parietal suture line and consequent physical separation of the two bones. An overall effect of this treatment is to cause the collapse of the bones on to the supporting grids. The addition of aspirin or indomethacin to the incubation medium containing collagenase causes inhibition of the release of ^{45}Ca, and although lysis of the connective tissue along the suture line

73

Table 1 Effect of collagenase on bone lysis determined in an assay using ^{45}Ca-labelled 4-d-old mouse calvaria as described previously[3]

Experiment	Concentration of collagenase (U ml^{-1})	Inhibitor	^{45}Ca release (test–control ratio)
a	5	—	2.34 ± 0.09
	10	—	2.58 ± 0.03
	15	—	3.31 ± 0.28
	15	Aspirin 10 µg ml	1.32 ± 0.04
	15	Indomethacin 1 µg ml^{-1}	1.46 ± 0.17
b	5 (Live bone)	—	2.39 ± 0.44
	5 (Dead bone)	—	1.15 ± 0.10
c	—	SC19220 25 µg ml^{-1}	0.64 ± 0.04
	15	—	2.72 ± 0.39
	15	SC19220 25 µg ml^{-1}	0.73 ± 0.09

Bones were cultured in Biggers medium containing 5% heat-inactivated rabbit serum (Wellcome Reagents). Collagenase was from *Clostridium histolyticum* (Sigma, type III). Aspirin was added to the medium from a stock solution of pure acetylsalicylic acid in Biggers, and indomethacin was added from a stock solution in ethanol to give a final concentration of 0.01% ethanol. This concentration had no significant effect on collagenase activity in these assays. The bones in experiment *b* were killed by immersion in distilled water for 24 h. SC19220 was provided by Searle Laboratories, Illinois. The results are the mean of four determinations ± s.d.m.

of the frontal and parietal bones continues, no macroscopically obvious breakdown of the calcified matrix occurs This treatment thus protects the bones from collapse other than at the suture line. No significant increase in ^{45}Ca release occurs if the bones are killed by immersion in distilled water before incubation with the enzyme, suggesting that an active cellular response is required, but severe weakening of the suture line is effected. Similar results are obtained using trypsin in place of collagenase.

Aspirin and indomethacin are inhibitors of prostaglandin synthesis[10] and some prostaglandins are known to stimulate resorption in live bone *in vitro*[11]. Although Brown and Pollock[14] found that 1 mg indomethacin ml^{-1} inhibited by 50% the action of collagenase on collagen as measured enzymatically, we found no such inhibition at drug concentrations below 100 µg ml^{-1} (Table 2). Similarly, in the bone assay system neither aspirin nor indomethacin prevented the separation of the bone along the suture lines. Our results suggest therefore that collagenase-induced osteolysis may be mediated by prostaglandins, an observation that is given support by the finding that SC19220, an antagonist of prostaglandin action[15], is able to inhibit the collagenase-induced release of ^{45}Ca (Table 1). Further evidence that prostaglandins probably mediate, in part or whole, the action of collagenase on bone is provided by results which show that increases in the prostaglandin E content of incubation media occur in a dose-dependent manner on addition of collagenase (Table 3).

Raisz *et al.*[16] have presented essentially similar evidence for prostaglandin mediation of complement-dependent stiumulation of bone resorption. It seems likely that a similar mechanism of resorption is shared by collagenase, trypsin

Table 2 Effect of indomethacin and aspirin on collagenase activity

Concentration of collagenase (U ml^{-1})	Inhibitor	Enzyme activity
0.5	—	0.35
2.0	—	0.76
5.0	—	1.00
5.0	Aspirin 5µg ml^{-1}	1.03
5.0	Aspirin 50 µg ml^{-1}	0.98
5.0	Indomethacin 1 µg ml^{-1}	0.93
5.0	Indomethacin 100 µg ml^{-1}	0.97

Enzyme activity is expressed as a ratio of the 5 U collagenase ml^{-1} value. Acid-soluble collagen was prepared as described previously[12]. The method of acetylation and of collagenase assay followed that of Ginslow and McBride[13]. The reaction mixture for the assay contained 2 mg ^{14}C-labelled collagen in a 1-ml volume, and the incubation time was 2 h at 37 °C. Aspirin was added from a stock solution of pure acetylsalicyclic acid in buffer. Indomethacin was added as a stock solution in ethanol and its effects were compared to 5.0 U ml^{-1} collagenase containing the same amount of ethanol.

and complement-sufficient heterologous serum, although stimulation of prostaglandin synthesis by the latter agent may depend on a different process. Inhibition of parathyroid hormone-stimulated bone resorption by aspirin[17] is not of the same order as that found on bone resorption stimulated

Table 3 Effect of collagenase on prostaglandin release by bone *in vitro*

Experiment	Prostaglandin released per bone (ng PGE$_2$ Eq)
Control	< 0.3
Collagenase 15 U ml^{-1}	25.6 ± 20.0
Collagenase 5 U ml^{-1}	8.6 ± 4.5
Collagenase 15 U ml^{-1} + indomethacin 1 µg ml^{-1}	< 0.3
Collagenase 15 U ml^{-1} + indomethacin 100 ng ml^{-1}	< 0.3

Unlabelled bones were used and the Biggers medium was supplemented with 0.5% bovine plamsa albumin (twice crystallised, Armour Pharmaceuticals) instead of rabbit serum to avoid complications with prostaglandins present in the serum. Indomethacin was added in ethanol to give a final concentration of 0.01%; the same amount of ethanol was added to all other cultures. For radioimmunoassay, a neutral lipid extract was followed by ethyl acetate at pH 4.7 (citric acid). The residue on evaporation of this was assayed using an antiserum (gift from May and Baker) raised against prostaglandin E$_2$ (PGE$_2$). Cross reaction was found to be 66% for PGE$_1$, but to be < 2% for PGs F$_{1\alpha}$, F$_{2\alpha}$, B$_2$, A$_1$ and A$_2$. The results are the mean of three determinations ± s.d.m.

by these agents. Indomethacin has been reported to have no significant inhibitory action on the stimulation by parathyroid hormone[16,18]. We have also found (unpublished results) that aspirin and indomethacin do not significantly inhibit bone resorption stimulated by vitamin A or 1,25-dihydroxycholecalciferol (provided by Professor I. MacIntyre). It is probable that the resorption achieved by these agents is not as dependent on prostaglandin synthesis for its perpetuation. Since collagenase is found to be released by bone during parathyroid hormone-stimulated resorption[19–21], however, prostaglandins may have some secondary function in physiologically normal osteolysis.

The digestion of bone matrix collagen by exogenous collagenase forms only one step in the highly coordinated process of bone resorption. To complete matrix breakdown, it is likely that other enzymes derived from the bone cells must act with collagenase extracellularly and that the resulting fragments must be endocytosed to undergo further breakdown. This is the same pattern of resorption as was suggested as taking place on parathyroid hormone or prostaglandin stimulation[22], neither of which is markedly affected by indomethacin. It therefore seems probable that the essential difference between the mechanisms of resorption of collagenase and parathyroid hormone, for example,

is in the manner in which stimulation is achieved. It seems reasonable to assume that it is the initial digestion of collagen which causes the bone cell response to exogenous collagenase. The collagen fragments may cause a bone cell memrane reactions which leads to prostaglandin synthesis, and the resulting increased prostaglandin levels may cause normal bone resorption to ensue. Similar membrane effects have been suggested as being involved in complement-sufficient serum-induced bone resorption[16].

Both prostaglandins and collagenase have been suggested as being responsible for the bone resorption associated with some tumours[5–9,23,24], inflamed gingival tissues[25] and dental cysts[26], and rheumatoid synovial tissue[27,28]. The work implicating prostaglandin release from these tissues indicates that indomethacin may be a suitable agent for inhibition of the resorption. Our results suggest that, although the lytic effect of these tissues might be partially dependent on their release of collagenase, indomethacin may still be equally effective in preventing bone destruction.

We thank Miss Tina Billingsley, Miss Diane Richardson and Miss Kathy Leane for assistance.

M. DOWSETT
A. R. EASTMAN
D. M. EASTY
G. C. EASTY
T. J. POWLES
A. M. NEVILLE

Unit of Human Cancer Biology, London Branch,
Ludwig Institute for Cancer Research,
  and Royal Marsden Hospital,
London SW3 6JB, UK

1 Beutner, E. H., Triftshauser, C., and Hazen, S. P., *Proc. Soc. exp. Biol. Med.,* **121**, 1082–1085 (1966).
2 Fullmer, H. M., and Gibson, W., *Nature,* **209**, 728–729 (1966).
3 Powles, T. J., Clark, S. A., Easty, D. M., Easty, G. C., and Neville, A. M., *Br. J. Cancer,* **28**, 316–321 (1973).
4 Powles, T. J., Dowsett, M., Easty, G. C., Easty, D. M., and Neville, A. M., *Lancet,* i, 608–610 (1976).
5 Dowsett, M., Easty, G. C., Powles, T. J., Easty, D. M., and Neville, A. M., *Prostaglandins,* **11**, 447–460 (1976).
6 Strauch, L., in *Tissue Interactions in Carcinogenesis* (edit by Tarin, D.), 399–433 (Academic, London, 1972).
7 Riley, W. B., Jr, and Peacock, E. E., Jr, *Proc. Soc. exp. Biol. Med.,* **124**, 207–210 (1967).
8 Yamanishi, Y., Maeyens, E., Dabbons, M. Kh., Ohyama, H., and Hashimoto, K., *Cancer Res.,* **33**, 2507–2512 (1973).
9 Dresden, M. H., Heilman, S. A., and Schmidt, J. D., *Cancer Res.,* **32**, 993–996 (1972).
10 Vane, J. R., *Nature new Biol.,* **231**, 232–235 (1971).
11 Klein, D. C., and Raisz, L. G., *Endocrinology,* **86**, 1436–1440 (1970).
12 Werb, Z., and Burleigh, M. C., *Biochem. J.,* **137**, 373–385 (1974).
13 Ginslow, M. T., and McBride, B. C., *Analyt. Biochem.,* **68**, 70–78 (1975).
14 Brown, J. H., and Pollock, S. H., *Proc. Soc. exp. Biol. Med.,* **135**, 792–795 (1970).
15 Sanner, J. H., *Archs intern. Med.,* **133**, 133–146 (1974).
16 Raisz, L. G., Sandberg, A. L., Goodson, J. M., Simmons, H. A., and Mergenhagen, S. E., *Science,* **185**, 789–791 (1974).
17 Powles, T. J., Easty, D. M., Easty, G. C., Bondy, P. K., and Neville, A. M., *Nature new Biol.,* **245**, 83–84 (1973).
18 Goldhaber, P., Rabadjija, L., Beyer, W. R., and Kornhauser, A., *J. Am. dent. Ass.,* **87**, 1027–1033 (1973).
19 Kaufman, E. J., Glimcher, M. J., Mechanic, G. L., and Goldhaber, P., *Proc. Soc. exp. Biol. Med.,* **120**, 632–637 (1965).
20 Fullmer, H. M., and Lazarus, G., *Israel J. med. Sci.,* **3**, 758–761 (1967).
21 Sakamoto, S., Sakamoto, M., Goldhaber, P., and Glimcher, M., *Biochem. biophys. Res. Commun.,* **63**, 172–178 (1975).
22 Dingle, J. T., *J. Bone Jt Surg.,* **55B**, 87–95 (1973).
23 Tashjian, A. H., Jr, Voelkel, E. F., Levine, L., and Goldhaber, P., *J. exp. Med.,* **136**, 1329–1343 (1972).
24 Voelkel, E. F., Tashjian, A. H., Jr, Franklin, R., Wasserman, E., and Levine, L., *Metabolism,* **24**, 973–986 (1975).
25 Goodson, J. M., Dewhirst, F. E., and Brunetti, A., *Prostaglandins,* **6**, 81–85 (1974).
26 Harris, M., Jenkins, M. V., Bennett, A., and Willis, M. R., *Nature,* **245**, 213–215 (1973).
27 Robinson, D. R., Tashjian, A. H., Jr, and Levine, L., *J. clin. Invest.,* **56**, 1181–1188 (1975).
28 Evanson, J. M., Jeffrey, J. J., and Krane, S. M., *Science,* **158**, 499–502 (1967).

Simple model for treating evolution of multigene families

THE mechanism by which coincidental or horizontal evolution occurs in the multigene families (such as ribosomal RNA genes) represents one of the most challenging problems of molecular evolution[1,2]. Smith[3] and Black and Gibson[4] put forward an elegant model of coincidental evolution, assuming homologous but unequal crossing over. They have shown, using computer simulations, that repeated unequal crossing over results in the gradual loss of gene lineages, and that eventually the whole family becomes fixed by a single gene lineage. Here the population dynamics of gene lineages within a multigene family is analogous to the population dynamics of mutant genes within a finite Mendelian population. In this communication I show that a quantitative theory of unequal crossover fixation can be developed based on the mathematical theory of population genetics[5,6], treating the behaviour of mutant alleles in finite populations.

Following Smith[3] and Black and Gibson[4], I consider intrachromosomal crossing over and concentrate on a single chromosomal lineage. Let us assume that there are n gene units within a family and that, initially, each unit is represented by a different gene lineage, so that there are n gene lineages at the start. To simplify the treatment, we assume that the mismatch between the two daughter chromosomes at the time of crossing over amounts to a shift of exactly one unit, so that one gene unit is either added or deleted through one unequal crossing over. We further assume that crossing over occurs at random along the n sites and the duplication and the deletion occur alternately, so that the total number of gene units is kept equal to the initial number, n. This corresponds to the action of natural selection holding the size of the gene family approximately constant. This model is illustrated in Fig. 1. Let us designate two consecutive crossovers (involving duplication and deletion) as one cycle of the process.

Now the problem is to analyse how gene lineages become fixed in the family or lost from it by continued unequal crossovers. This problem is quite similar to that of fixation or loss of mutant alleles in finite populations; the unequal crossover corresponds to sampling of genes at reproduction. Therefore, we can treat the frequency change of gene lineages in a multigene family as we treat the gene frequency change in standard population genetics, and in particular we can apply the diffusion model developed by Kimura[7].

The frequency of a particular lineage in the family after a certain number of cycles of unequal crossovers is denoted by x. Note that, at the start, $x = 1/n$, because each lineage is singly represented in the family at the start. To apply the diffusion equation method in population genetics to our problem, we need to evaluate the mean ($M_{\Delta x}$) and the variance ($V_{\Delta x}$) of the frequency change per unit time (that is per cycle of unequal crossovers in this case). These quantities can be computed using Table 1. From the last two columns of Table 1, we can obtain $M_{\Delta x}$ and $V_{\Delta x}$.

$$M_{\Delta x} = \frac{1}{n}x(1-x) - \frac{1}{n}x(1-x) = 0 \tag{1}$$

$$V_{\Delta x} = \frac{1}{n^2}x(1-x) + \frac{1}{n^2}x(1-x) = \frac{2}{n^2}x(1-x)$$

Table 1 was constructed from the consideration that duplication or deletion by unequal crossover occurs at random on the chromosome, so that the probability that a particular lineage increases its size by one unit by duplication, or decreases its size by one unit by deletion, is in each case equal to frequency.

Once the mean and the variance of the frequency change per cycle are obtained, it is possible to apply Kimura's theory of gene frequency change. Note that the zero mean change ($M_{\Delta x} = 0$) in formula (1) implies that the population dynamics for selectively neutral mutants can be applied. In other words, my model is a diffusion process with a zero mean change and a non-zero variance as given in formula (1) and each gene lineage has equal chance of eventually becoming fixed in the family. Now, the dynamics of mutant frequency change in a finite population is a diffusion process with a mean $M_{\Delta x} = 0$ and a variance $V_{\Delta x} = x(1-x)/2N_e$ in case of neutral genes and in a population with effective size N_e. On the other hand, the
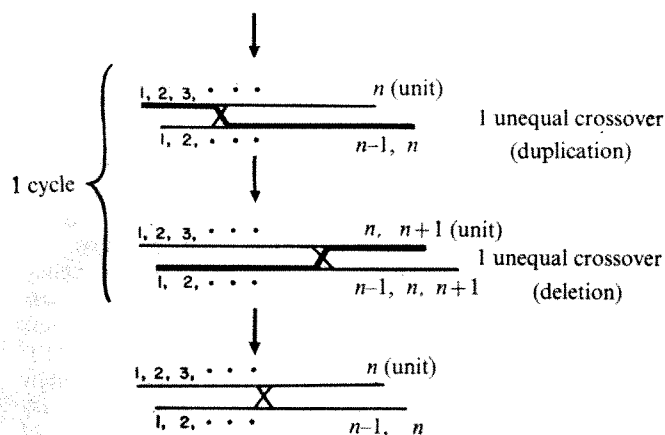
Fig. 1 Diagram illustrating the model.

crossover fixation time of one of the n lineages is equivalent to the time until fixation of a mutant allele excluding the cases of loss in population genetics. Kimura and Ohta[8] have obtained the mean fixation time of a selectively neutral mutant in terms of the number of generations (excluding loss). Therefore by analogy, or by replacing N_e with $n^2/4$ in Kimura and Ohta formula, we get the mean of the crossover fixation time in terms of the number of cycles in the following form

$$t_1(p) = -\frac{1}{p}\{n^2(1-p)\log_e(1-p)\} \qquad (2)$$

where p is the initial frequency of the gene lineage and equal to $1/n$. When n is large or p is small, this formula becomes approximately

$$t_1(p) \approx n^2 \qquad (3)$$

Kimura and Ohta also obtained the variance of the time until fixation of the neutral mutant in finite populations[9]. Again, by replacing N_e with $n^2/4$ in their formula, we get

$$\text{var}(t_1) \approx 0.286n^4 \qquad (4)$$

I have checked the above results by computer simulation and the agreements between the predicted and the observed values are generally satisfactory, both for the mean and the variance of crossover fixation time.

Pursuing the analogy with population genetics, it may be convenient to define a quantity similar to homozygosity. In fact, Smith[3] used the quantity called the clonality in his simulation studies, which is equal to the sum of squares of frequencies of gene lineages, and which is equivalent to the homozygosity;

$$C = \sum_i x_i^2 \qquad (5)$$

where the subscript i indicates the ith gene lineage. The expected change of clonality per one cycle of unequal crossovers can be calculated again analogously with that of homozygosity in finite populations.

$$E\{\Delta C\} = E\{\sum_i (x_i+\Delta x_i)^2 - \sum_i x_i^2\} = E\{\sum_i (\Delta x_i)^2\}$$

$$= (1-C)\frac{2}{n^2} \qquad (6)$$

where E stands for taking the expectation. This formula can be derived by using the formula for $V\Delta_x$ in (1) taking all lineages (that is, all i) of the family into account. Again, $n^2/2$ corresponds to $2N_e$ for the change of homozygosity in finite populations.

The present model may be an oversimplification of the actual process of crossover fixation; in particular it assumes

that the duplication and the deletion occurs by one unit. Smith[3] has considered the cases where the duplication and deletion are allowed to vary by more than one unit and his simulation results suggest that this latitude accelerates the diffusion process. Although the exact theoretical treatment of such cases is difficult, very rough estimation can be done as follows. If the mean number of gene units duplicated or deleted by one unequal crossover is m, then one crossover should correspond to $m/2$ cycles of my model, provided that the arrangement of gene lineages is more or less random along the chromosome. By comparing Smith's results (Fig. 3 in ref. 3) with this estimation, it can be predicted that the random arrangement of gene lineages may be obtained when $m \approx n/10$ or when the allowed latitude (which is approximately equal to m) is roughly 10% of the total number of gene units. This

Table 1 Calculations for derivation of equation (1)

Duplication phase		Deletion phase		Total (1 cycle)	
Change of x	Probability	Change of x	Probability	Change of x	Probability
$+\frac{1}{n}$	x	$-\frac{1}{n}$	x	0	x^2
		0	$1-x$	$+\frac{1}{n}$	$x(1-x)$
0	$1-x$	$-\frac{1}{n}$	x	$-\frac{1}{n}$	$x(1-x)$
		0	$1-x$	0	$(1-x)^2$

follows because the observed crossover fixation time in this case ($\approx 2,000$) corresponds to that predicted by the present model, that is $2/10 \times 10,000$. It should be noted here that Smith stopped his simulations at the clonality of 0.9 or when the frequency of a lineage reached about 0.95. The time required to reach this frequency is slightly smaller but not much less than the complete fixation time[10].

I shall compare Smith's results with my prediction in more detail. He obtained the crossover fixation time of very roughly 8,000 for the case of latitude $(m) = 1$ and roughly 2,000 for $m = 5 \sim 25$. The predicted values by my model become for these cases: 20,000, 4,000, 2,000 and 800 crossovers corresponding to the four values of m (1, 5, 10 and 25). This would seem to imply that, for cases when $m < n/10$, the gene lineages are more clustered than random arrangement and hence the actual crossover fixation time is smaller than the predicted values. This is because the variance of the frequency change ($V_{\Delta x}$) gets larger by the correlated changes in such circumstances. If m is larger ($m > n/10$), the gene lineages become more dispersed than random arrangement and the theory gives an underestimation.

The analyses in this communication are not complete, but they at least provide a theoretical basis for further analysis in quantitative terms of the mechanism of coincidental evolution of multigene families.

I thank Dr Motoo Kimura for discussions and suggestions.

TOMOKO OHTA

National Institute of Genetics,
Mishima, Shizuoka-ken,
411, Japan

[1] Hood, L. E., Wilson, J. H., and Wood, W. B., *Molecular Biology of Eucaryotic Cells* (Benjamin, Menlo Park, 1975).

[2] Hood, L., Campbell, J. H., and Elgin, S. C. R., *A. Rev. Genet.*, 9, 305–353 (1975).
[3] Smith, G. P., *Cold Spring Harb. Symp. quant. Biol.*, 38, 507–513 (1974).
[4] Black, J. A., and Gibson, D., *Nature*, 250, 327–328 (1974).
[5] Crow, J. F., and Kimura, M., *An Introduction to Population Genetics Theory* (Harper and Row, New York, 1970).
[6] Kimura, M., and Ohta, T., *Theoretical Aspects of Population Genetics* (Princeton University Press, Princeton, 1971).
[7] Kimura, M., *J. appl. Probability*, 1, 177–232 (1964).
[8] Kimura, M., and Ohta, T., *Genetics*, 61, 763–771 (1969).
[9] Kimura, M., and Ohta, T., *Genetics*, 63, 701–709 (1969).
[10] Maruyama, T., and Kimura, M., *Jap. J. Genet.*, 46, 407–410 (1971).

Do increases in enzyme activities during muscle differentiation reflect expression of new genes?

TRANSITIONS from embryonic isoenzymic forms to mature adult forms are known to occur during *in vitro* myogenesis for creatine phosphokinase[1] (CPK) and other muscle enzymes[2]. More recently, similar transitions have been shown to occur for myosin[3,4], and actin[5,6]. Both CPK activity[7,8] and amino acid incorporation into myosin[9–11] increase during the early stages of differentiation *in vitro* when cells fuse to form multinucleate myotubes, and it is tempting to assume that these increases require the production of adult form mRNA after activation of a "battery" of genes for myogenesis[2]. Yaffe and Dym[11] have proposed a model of myogenesis in which mRNAs for characteristic muscle proteins are synthesised before cell fusion and stored in a "masked" form until activated by the fusion process, thus giving rise to simultaneous increases in muscle-specific protein synthesis and cell fusion. We describe here experiments which show large increases in CPK activity at the time of early cell fusion which cannot result from activation of "masked" muscle-specific mRNAs and which do not require adult form mRNAs at all. Using a new technique which allows measurement of low levels of CPK isoenzymes, we have shown that the early increase in CPK activity during the first 2 d of culture is of the embryonic, or brain, form, and only later does the adult muscle form make a major contribution to the total increase in enzyme activity.

Embryonic chick skeletal muscle cells were prepared and grown as described previously[12]. Freeze-thawed cell pellets were extracted with 10% sucrose and samples from the centrifuged extract were assayed for CPK activity and protein as described previously[8] to determine specific activity. Isoenzymes were separated by discontinuous polyacrylamide gel electrophoresis[13] and the total specific activity was distributed among the three forms according to the ratios of the peaks of activity obtained after slicing the gel. Appropriate control experiments show that recovery of activity from the gel slices is independent of enzyme concentration and is the same for muscle and brain forms.

Fig. 1 CPK isoenzyme patterns of extracts from 74 h (●) and 96 h (○) muscle cell cultures. A sample (50 µl) of extract was subjected to electrophoresis at 4 °C on 10% polyacrylamide gels with 5% cross linkage[12] for 3 h at 3 mA per tube. Slices (1 mm) of the frozen gels were placed in standard assay mixture and assayed for CPK activity as described previously[8]. CPK specific activity is expressed as international units (1 IU converts 1 µmol of creatine phosphate to creatine per min of incubation at 37 °C) per mg soluble protein in the original extract.

Fig. 2 Specific activities of the three isoenzymes of CPK (BB, ■; MB, ○; and MM, ●) during the first 2 d of culture. Cell fusion (□, broken line) was determined as previously described[8]. Each point represents a determination from a single gel and the combined data of four separate cell preparations are shown.

Creatine phosphokinase is a dimer of brain-type (B) or muscle-type (M) subunits and skeletal muscle differentiation is accompanied by a transition from BB isoenzyme, through the intermediate MB form, to the mature MM form. Figure 1 illustrates the resolution of the three forms on polyacrylamide gels and the continuing transition between 72 h and 96 h of culture. These results are consistent with those obtained by qualitative staining methods involving fluorescence[1] or colour[2] changes associated with enzyme-coupled reactions. Until now, however, no method has been sufficiently sensitive to study isoenzyme changes during earlier stages of culture when increases in cell fusion and enzyme specific activity are just beginning.

Between 17 h and 42 h of culture, fusion increases from

1 to 20% (Fig. 2) and total CPK specific activity increases sixfold. As Fig. 2 shows, this increase is due almost entirely to the BB embryonic form of the enzyme, the mature and intermediate forms showing little change. Whether the simultaneous increase in cell fusion and BB–CPK activity is anything other than coincidental remains to be determined. Although we, and many others, have shown that apparently normal increases in CPK activity can occur at later stages in the absence of cell fusion[12], this has not yet been established for the early increase described here. Earlier[8], we described a transient increase in CPK activity occurring at the time of the initial increase in percentage fusion and before the much larger increases in enzyme activity which occurred later. It is possible that that transient increase is related to the increase in the BB form described here, although direct comparison is not possible since we now use cells prepared differently which fuse and develop much earlier. In this respect, it is interesting that muscle-forming cultures of mouse teratocarcinoma cells show an increase in CPK at the time of myotube formation, followed by decrease, but without any isoenzyme transition from the BB form[14].

The assumption that muscle-specific forms are also involved underlies experiments in which CPK activity and myosin synthesis were "superinduced" by actinomycin D during and shortly before cell fusion. These experiments have been used to support the model involving prefusion synthesis of "masked" mRNAs for protein characteristic of the differentiated state[11]. We have been able to obtain stimulations of 30–50% of total CPK specific activity (and of total activity per culture) by exposing either 22-h or 42-h cultures to actinomycin D (Sigma, 1 μg ml⁻¹) for 4 h. Although an increase of this size specific for the MM or MB forms would be readily detected by our electrophoretic method at these stages, we have found no significant change in the isoenzyme pattern, suggesting that the stimulation is either nonspecific or involves only the BB form. It is fair to say, however, that Yaffe and Dym[11] used rat cells, rather than chick, in rather different culture conditions and obtained larger stimulations by actinomycin D, but, at the very least, our results argue forcibly for the need to demonstrate the isoenzyme transition before conclusions in terms of stored mRNAs for muscle-specific molecules can be made.

In general, we would advise caution in interpreting data on increases in enzyme activities or synthesis during differentiation in terms of new gene expression until they are shown to be due to the authentic mature form of the protein. In view of reports of isoenzyme-like transitions in myosin heavy chains[3,4], these considerations are also applicable to the interpretation of early increases in myosin synthesis during differentiation[10,11], and may influence conclusions drawn from studies of myosin mRNA during myogenesis, since the myosin mRNA synthesis observed in early cultures[9,15] may be of the form that is not muscle specific.

This work was supported by the Muscular Dystrophy Group of Great Britain and the Medical Research Council.

GLENN E. MORRIS
MELANIE PIPER
ROBIN COLE

School of Biological Sciences,
University of Sussex,
Falmer, Brighton BN1 9QG, UK

1 Morris, G. E., Cooke, A., and Cole, R. J., *Expl Cell Res.*, 74, 582–584 (1972).
2 Turner, D. C., Maier, V., and Eppenberger, H. M., *Devl Biol.*, 37, 63–89 (1974).
3 Sreter, F. A., Balint, M., and Gergely, J., *Devl Biol.*, 46, 317–325 (1975).
4 Chi, J. C., Fellini, A., and Holtzer, H., *Proc. natn. Acad. Sci. U.S.A.*, 72, 4999–5003 (1976).
5 Whalen, R. G., Butler-Browne, G. S., and Gros, F., *Proc. natn. Acad. Sci. U.S.A.*, 73, 2018–2022 (1976).
6 Storti, R., Coen, D., and Rich, A., *J. Cell Biol.*, 67, 420a (1975).
7 Shainberg, A., Yagil, G., and Yaffe, D., *Devl Biol.*, 25, 1–29 (1971).
8 Morris, G. E., and Cole, R. J., *Expl Cell Res.*, 75, 191–199 (1972).
9 Morris, G. E., *et al.*, *Cold Spring Harb. Symp. quant. Biol.*, 37, 535–541 (1972).
10 Paterson, B., and Strohman, R. C., *Devl Biol.*, 29, 113–138 (1972).
11 Yaffe, D., and Dym, H., *Cold Spring Harb. Symp. quant. Biol.*, 37, 543–547 (1972).
12 Morris, G. E., Piper, M., and Cole, R. J., *Expl Cell Res.*, 99, 106–114 (1976).
13 Davis, B. J., *Ann. N.Y. Acad. Sci.*, 121, 404–427 (1964).
14 Gearhart, J. D., and Mintz, B., *Cell*, 6, 61–66 (1975).
15 Buckingham, M. E., Caput, D., Cohen, A., Whalen, R. G., and Gros, F., *Proc. natn. Acad. Sci. U.S.A.*, 71, 1466–1470 (1974).

Inhibitory effect of α-(1→6)-heterogalactan on oocyte maturation of starfish induced by 1-methyladenine

THE surface of the starfish oocyte is thought to carry receptors for 1-methyladenine (1-MA)[1,2], which induces maturation division by some mechanism which is at present little understood. Using α-(1→6)-heterogalactan (HGaln) and α-(1→6)-galacto-oligosaccharides, which strongly inhibit the induction of maturation division by 1-MA, we have now obtained evidence that the receptor for HGaln is on the cell surface. We found that in the presence of HGaln, 1-MA did not activate the cell surface.

HGaln is extracted from the fruiting bodies of *Lentinus edodes* and has a main chain of (1→6)-linked α-D-galacto-pyranose residues, parts of which are substituted in the 2-position with either single L-fucopyranose or D-mannopyranose residues[3]. HGaln gives one peak only on ultracentrifugation (S = 1.55), zone electrophoresis and gel filtration.
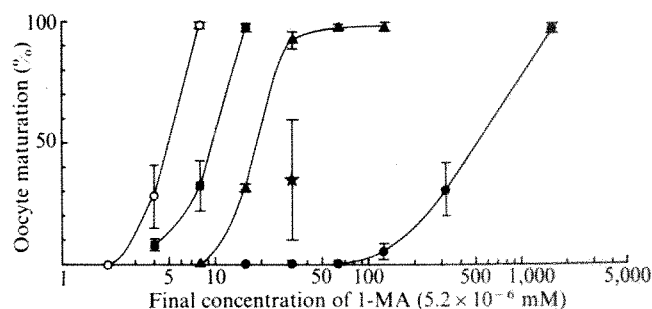
Fig. 1 Dose–response curves of the starfish oocytes to 1-MA in the presence of α-(1→6)-HGaln(FC, 0.83%) (●), or α-(1→6)-galactotetraose(FC, 0.83%) (▲), or in the absence of any sugars (○). ■, Dose–response of the washed oocytes by HASW after incubation with 0.83% α-(1→6)-HGaln for 30 min at 0–4 °C. Note the inhibitory effect of α-(1→6)-HGaln and α-(1→6)-galactotetraose on the oocyte maturation. At the star 192 × (5.2 × 10⁻⁶ mM) of 1-MA was mixed with 5% α-(1→6)-HGaln at the same volume for about 1 h at 0–4 °C, and 10 μl of the mixture was added to 20 μl of the starfish oocyte suspension in HASW.

Specimens of the starfish *Asterina pectinifera* were collected from Muro Beach, Japan, and kept in an aquarium at 0–4 °C. Ovaries were removed with forceps in calcium-free Hervst's artificial seawater (HASW), and oocytes were prepared as before[4]. After incubation for 1 h in HASW at 0–4 °C, supernatant was decanted, and the oocytes were resuspended in HASW. This washing procedure was repeated to remove follicle cells. Although oocytes stored at 0–4 °C in HASW maintained their sensitivity to 1-MA at the same level for several days, it was necessary to test this sensitivity before each experiment in order to obtain quantitative results. Serial experiments were
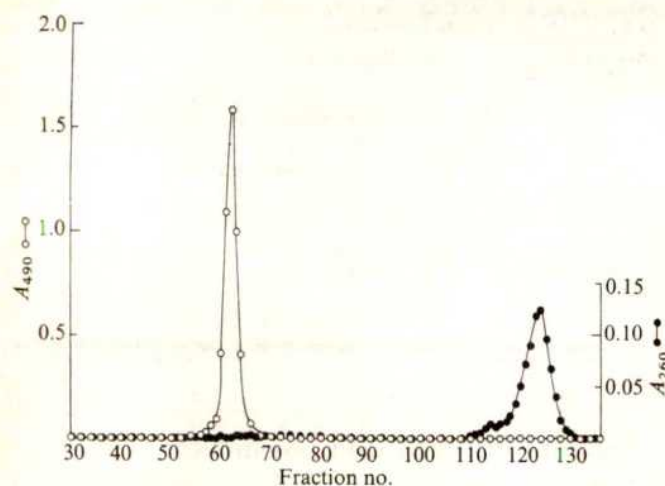
Fig. 2 Chromatography of the mixture of α-(1→6)-galacto-tetraose and 1-MA on Sephadex G-25. The column (2.64 × 85 cm) was washed with 500 ml of redistilled water and re-equilibrated with 1 l of 2/3 HASW. The applied sample contained 0.4 ml of 5% α-(1→6)-galactotetraose aqueous solution, 0.4 ml of 1 mM 1-MA aqueous solution and 0.8 ml of HASW. Elution was carried out with 2/3 HASW. Fractions (5 ml) were collected at a flow rate of 10 ml h^{-1}. ○, A_{260}; ●, carbohydrate measured by the phenol-sulphuric acid method at 490 nm.

performed with the oocytes of one animal. To detect the response of the starfish oocytes to 1-MA, 10 µl of 1-MA dissolved in redistilled water was added from a microcapillary pipette into 20 µl of oocyte suspension dispersed in HASW on

Table 1 Effect of the polysaccharides or monosaccharides that constitute α-(1→6)-HGaln, or oligosaccharides from the partial hydrolysate, on oocyte maturation induced by 1-MA in *Asterina pectinifera*

Test samples	Oocyte maturation (%) mean ± s.e.
(A)	
α-(1→6)-Mannan	100 ± 0
Glycogen from rabbit liver	97.0 ± 1.3
Dextran T 40 (MW, 40,000)	98.0 ± 1.8
(B)	
D-Galactose	100 ± 0
D-Mannose	95.8 ± 3.4
L-Fucose	95.7 ± 2.3
(C)	
α-(1→6)-Heterogalactan	0 ± 0
α-(1→6)-Galactobiose	3.4 ± 1.6
α-(1→6)-Galactotriose	1.9 ± 1.6
α-(1→6)-Galactotetraose	0 ± 0
(D)	
Maltose $\underset{\alpha}{\overset{1\ 4}{G\longrightarrow G}}$	100 ± 0
Maltotriose $\underset{\sigma\quad\alpha}{\overset{1\ 4\ 1\ 4}{G\longrightarrow G\longrightarrow G}}$	98.9 ± 0.9
Maltotetraose $\underset{\alpha\quad\alpha\quad\alpha}{\overset{1\ 4\ 1\ 4\ 1\ 4}{G\longrightarrow G\longrightarrow G\longrightarrow G}}$	100 ± 0
Isomaltose $\underset{\alpha}{\overset{1\ 6}{G\longrightarrow G}}$	100 ± 0
Isomaltotriose $\underset{\alpha\quad\alpha}{\overset{1\ 6\ 1\ 6}{G\longrightarrow G\longrightarrow G}}$	100 ± 0

In each experiment, the minimum concentration of 1-MA for inducing 100% oocyte maturation was added. This was the most suitable concentration for detecting the inhibitory effect. Concentration of test samples was 0.83%.

α-(1→6)-Mannan was provided by Dr T. Nakajima, Tohoku University.

G, Glucose; MW, molecular weight.

hole glasses (well size, 22 × 2 mm). The mixture was agitated thoroughly with a fine needle. Whole oocytes and those with broken germinal vesicles (OBGV) that had been induced spontaneously were counted under the microscope within 10 min. Breakdown of germinal vesicles began after 15 min of incubation and ended within 30 min of addition of 1-MA, independent of concentration. Therefore the hole glasses were covered with glass plates and after 30–40 min total OBGV was recorded. Each experiment was carried out three times and the percentage of oocyte maturation was calculated as OBGV induced when 1-MA was added to total oocytes. To test the inhibitory effect, 10 µl of the test solution dissolved in redistilled water was added to 40 µl of oocyte suspension in HASW and agitated. A sample of 10 µl of 1-MA solution in redistilled water was then added within 10 min.

As Fig. 1 shows, these oocytes underwent 100% maturation when 1-MA was added in a final concentration (FC) of more than $8 \times (5.2 \times 10^{-6}$ mM). On the other hand, maturation was completely inhibited by 0.83% (FC) heterogalactan in the presence of the same concentration of 1-MA (FC, $8 \times (5.2 \times 10^{-6}$ mM)). FC of $1,500 \times (5.2 \times 10^{-6}$ mM) 1-MA was needed to induce 100% maturation in the presence of 0.83% α-(1 → 6)-HGaln.

The question is whether the inhibitory effect of HGaln was due to D-galactose, or D-mannose. It was also necessary to exclude nonspecific absorption of polysaccharides on the cell surface. We therefore tested for inhibition by other poly-saccharides in experiments summarised in Table 1.

No inhibitory effect was detected with α-(1→6)-mannan isolated from the mutant *Saccharomyces cerevisiae* X2180 (ref. 5), Dextran T 40 (molecular weight 40,000, Pharmacia), or glycogen from rabbit liver (Sigma). This excludes nonspecific inhibition by polysaccharides of the induction of oocyte maturation by 1-MA. D-galactose, L-fucose and D-mannose

Fig. 3 Electron micrographs of the oocytes of *Asterina pectinifera*. *a*, Normal cell surface of the oocytes. The surface is covered with the continuous vitelline envelope (VE). CG, Cortical granule (×22,050). *b*, Surface of the trypsinised oocyte. The vitelline envelope is almost removed by trypsinisation for 4 h at 24 °C with 1% trypsin solution (Difco, 1:250) dissolved in HASW (×22,050).
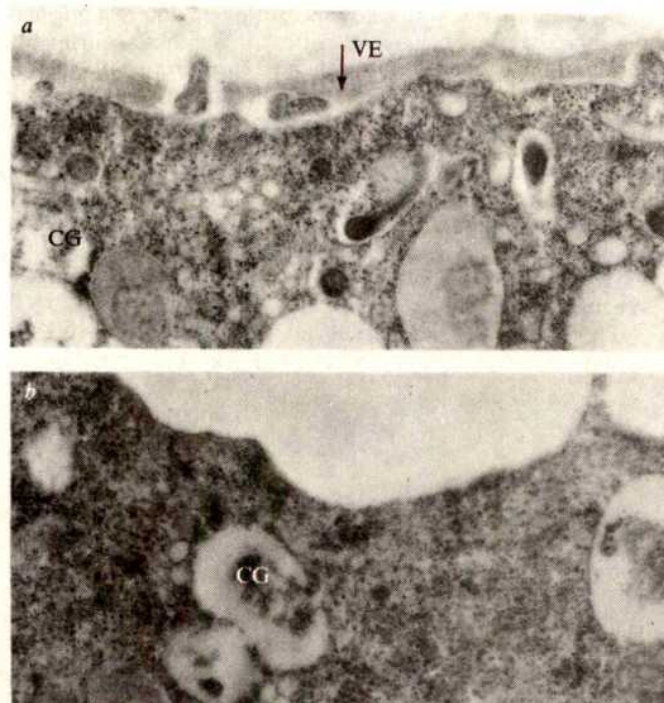
were also ineffective at the same concentration. α-(1→6)-Galactobiose, α-(1→6)-galactotriose and α-(1→6)-galacto-tetraose, obtained from the partial hydrolysate of α-(1→6)-HGaln[3], were inhibitory (Table 1). The minimum concentration of 1-MA that induced 100% maturation was $30 \times (5.2 \times 10^{-6}$ mM) in the presence of 0.83%(FC) α-(1→6)-galacto-tetraose (Fig. 1). No gluco-oligosaccharides had any inhibitory effect, even if the binding form was α-(1→6)-linked types as the galactosides. Thus the inhibition was due to phenomena specific to α-(1→6)-galactan or α-(1→6)-galacto-oligosaccharides.

We then considered whether the inhibitor might bind directly to 1-MA in the extracellular space, or in the cytoplasm after penetrating the membrane, or on the cell surface. The first possibility was excluded because when $192 \times (5.2 \times 10^{-6}$ mM) 1-MA was mixed with an equal volume of 5% α-(1→6)-HGaln for about 1 h at 0–4 °C, and 10 μl of this mixture was added to 20 μl of starfish oocyte suspension in HASW, the extent of inhibition was 64.4% compared with 100% with the same concentration of 1-MA (FC, $32 \times (5.2 \times 10^{-6}$ mM)) and α-(1→6)-HGaln (FC, 0.83%) after preincubation of oocytes with α-(1→6)-HGaln.

This was supported by chromatography, using α-(1→6)-galactotetraose because α-(1→6)-HGaln showed trace absorption at 260 nm. A mixture of 0.4 ml of 5% α-(1→6)-galacto-tetraose, 0.4 ml of 1-MA solution and 0.8 ml of HASW was stored at 4 °C for 3 h and adapted for a Sephadex G-25 column $(2.64 \times 85$ cm). Elution was carried out with a 2/3 concentration of HASW. Fractions (5 ml) were collected at a flow rate of 10 ml h^{-1}. 1-MA and α-(1→6)-HGaln were completely separated by the chromatography (Fig. 2).
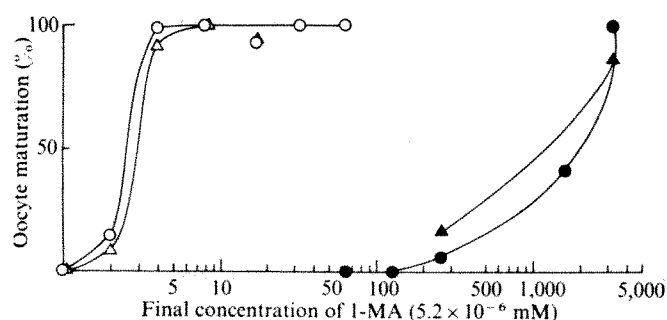
Fig. 4 Dose–response curve of the trypsinised oocytes to 1-MA and α-(1→6)-HGaln. The trypsinisation was carried with 1% trypsin solution for 4 h at 24 °C. ○, Untreated oocyte, no α-(1→6)-HGaln; ●, untreated oocyte, +α-(1→6)-HGaln; △, trypsinised oocyte, no α-(1→6)-HGaln; ▲, trypsinised oocyte, +α-(1→6)-HGaln. The trypsinised oocytes showed the same degree of response for 1-MA and α-(1→6)-HGaln as the untreated oocytes.

As the S value of α-(1→6)-HGaln is 1.55, it is unlikely that it penetrates the cell membrane. This was confirmed as follows. After pretreatment of oocytes with 0.83% α-(1→6)-HGaln for 30 min at 0–4 °C, they were washed twice with HASW, and their sensitivity to 1-MA was examined. The minimum concentration that induced 100% maturation was about $15 \times (5.2 \times 10^{-6}$ mM) (Fig. 1), far less than the value obtained in the presence of exogeneous α-(1→6)-HGaln. Some of the inhibitory effect, however, remained, even after washing. This suggests that the receptor of α-(1→6)-HGaln is present on the cell surface.

To locate the receptors of α-(1→6)-HGaln and 1-MA, the vitelline envelope was almost removed from oocytes by incubation for 4 h at 24 °C with 1% trypsin dissolved in HASW (Fig. 3b). Then oocytes showed the same degree of sensitivity to 1-MA and the same degree of inhibition as untreated oocytes (Fig. 4). The receptor of 1-MA and α-(1→6)-HGaln may therefore be on the plasmalemma.

Department of Anatomy,
Tohoku University School of Medicine,

MARIKO SHIDA

Department of Agricultural Chemistry,
Faculty of Agriculture, Tohoku University,
Sendai, Japan

1 Kanatani, H., and Hiramoto, Y., *Expl Cell Res.*, **61**, 280–284 (1970).
2 Dorée, M., and Guerrier, P., *Expl Cell Res.*, **96**, 296–300 (1975).
3 Shida, M., Haryu, K., and Matsuda, K., *Carbohyd. Res.*, **41**, 211–218 (1975).
4 Shida, H., Hirai, S., and Shida, M., *Expl Cell Res.*, **73**, 509–512 (1972).
5 Raschke, W. C., Kern, K. A., Antalis, C., and Ballou, C. E., *J. biol. Chem.*, **248**, 4660–4666 (1973).

Production of 14C- and 11C-labelled biomolecules using ionised gases

To establish a rapid method for eventual use in incorporating 11C into important biomolecules, we have studied the interaction of ionised gases with a number of natural products. Our findings on the degradation of cholesterol induced by ionised gases[1] led us to anticipate that such labelling might occur if the gas contained a radioactive label and we present results with 14CO which verify the feasibility of the method. The same approach could also be used with 11C, a positron-emitter which is too short lived for most conventional organic chemical labelling procedures[2]. The resulting tracer biomolecules should have widespread applications in chemistry, biochemistry, biology and nuclear medicine.

The methods and apparatus are broadly similar to those already described[1]. In brief, the principle is as follows (Fig. 1): 14CO gas is introduced into a vacuum system where it is bombarded by a 190-eV electron beam (E) to yield[3] 14CO+, 14C+, O+, atomic 14C and O, and excited 14CO.

Fig. 1 Principle of the labelling method. Electrons emitted by a filament, F, are collimated by a magnetic field, B, to form an axial electron beam E, which is accelerated towards the cylinder C by the potential V_c and collected on a plate P on further acceleration by the potential V_p. Inside the cylinder C, the electrons collide with 14CO molecules (O) and produce an ionised gas containing atoms, ions and excited molecules. These highly reactive species drift to the inner surface of the cylinder C which is coated with a thin organic film. The reactions occurring on the surface produce 14C-labelled compounds which are analysed by chromatography and crystallisation to constant specific activity.
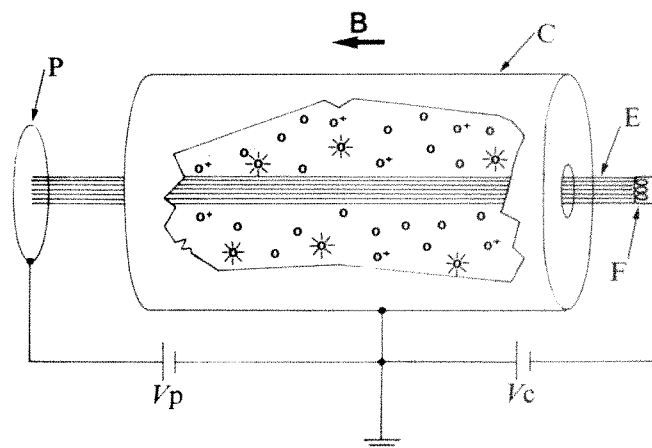
Fig. 2 Tracing of an autoradiogram after thin-layer chromatography of the ^{14}C products obtained from the interaction of organic substrates with ^{14}CO ionised gas. *A*, The position of the parent compound which was added as a marker is indicated by an arrow in each case. The starting materials were: *a*, oestradiol; *b*, 5,6-dimethyl-benzimidazole; *c*, D-arabinose; *d*, thymine; *e*, glycine. Chromatographic systems were: *a*, benzene–ethyl acetate (3:2); *b*, *c*, methanol–chloroform (4:1); *e*, methanol–benzene (2:3), all on silica gel; and *d*, *n*-butanol–acetic acid–water (80:12:30), on cellulose. *B*, Two dimensional analysis of cholesterol after exposure to ^{14}CO ionised gas. Chromatographic solvents were: I, hexane–ethyl acetate (3:2) and II, chloroform–acetone (95:5). Most of the total activity was recovered in spots 2, 4 and 6. By placing reference compounds with the labelled material before analysis, spots 2 and 4 were found to coincide with cholesterol and cholesterol formate, respectively. The identity of the ^{14}C-cholesterol fraction was confirmed by crystallisation to constant specific activity.
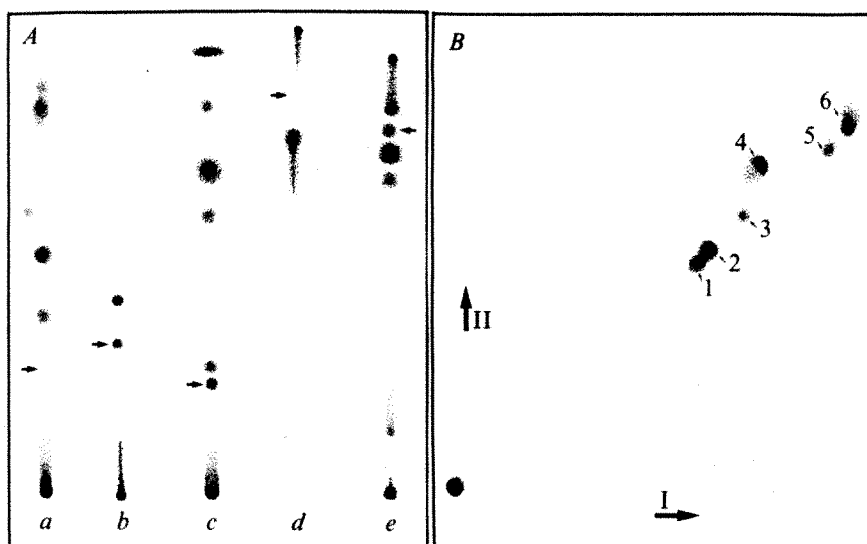
These species drift to the inner surface of the cylinder (C) which is coated with organic material. Reactions with the substrate yield a number of labelled products, as shown by chromatographic analysis (Fig. 2*A*). Steroids, proteins, sugars and a number of heterocyclic compounds were treated thus and gave radioactive products in varying yields, some of which migrated with the starting material. Typically, a 2-min bombardment of 10–100 μg of substrate with carrier-free ^{14}CO gave a total yield of labelled products of up to 1 mCi mmol^{-1}. Figure 2*A* shows autoradiograms obtained from a one-dimensional thin-layer chromatographic analysis of experiments with oestradiol, thymine, 5,6-dimethylbenzimidazole, D-arabinose and glycine. A similar analysis was carried out in two dimensions for cholesterol (Fig. 2*B*) and reveals the presence of six major labelled compounds. Two of these (2 and 4) migrate at the positions of cholesterol and cholesterol formate, respectively. The identity of compound 2 (Fig. 2*B*) as the labelled, but otherwise unaltered, parent molecule was established by crystallisation to constant specific activity of the free sterol and of its acetate derivative (Table 1). ^{14}C-cholesterol yields of up to 0.1 mCi mmol^{-1} were demonstrated in this manner. Although this value may seem low compared with chemical

labelling procedures, the replacement of ^{14}C by ^{11}C with its 10^8-fold shorter lifetime, would increase the activity to as much as 10 Ci μmol^{-1}. Even higher yields can be obtained for some of the secondary products of other substrates. This method thus seems particularly promising for the application of ^{11}C in bioorganic sciences and nuclear medicine.

L. SANCHE
J. E. VAN LIER

Department of Nuclear Medicine and Radiobiology,
University of Sherbrooke Medical Center,
Sherbrooke, Quebec, Canada J1H 5N4

1 Sanche, L., and van Lier, J. E., *Chem. Phys. Lipids* 16, 225 (1976).
2 Massey, H. S. W., *Electronic and Ionic Impact Phenomena*, II (Clarendon, Oxford, 1969).
3 Clark, J. D., and Buckingham, P. D., *Short-lived Radioactive Gases for Clinical Use*, 215–256 (Butterworths, London, 1975).

Erratum

In the article "Antiviral and cell growth inhibitory activities reside in the same glycoprotein of human fibroblast interferon" by E. Knight, Jr (*Nature*, 262, 302; 1976) the first line of column 1, page 303 should read . . . was exercised when observing growth inhibition with im-. . . . and not as printed.

Table 1 Crystallisation of ^{14}C-cholesterol from the interaction of cholesterol with an ionised ^{14}CO gas

Crystallisation		Specific activity (d.p.m. g^{-1})
No.	Solvent	
1	Methanol	13,840
2	Methanol	9,964
3	Methanol	9,341
4	Methanol	9,022
5	Petroleum ether–ethyl ether	8,064
6	Petroleum ether–ethyl ether	7,983
7	Petroleum ether–ethyl ether	8,075
8	Petroleum ether–ethyl ether	7,915
9	Methanol–ethyl ether (3-acetate)	8,131
10	Methanol–ethyl ether (3-acetate)	7,974
11	Methanol–ethyl ether (3-acetate)	8,038

Cholesterol (60 μg) was bombarded as described in the text. Labelled products were recovered from the cylinder C (Fig. 1) with ethyl acetate and diluted with 4.87 g of carrier cholesterol. After the eighth crystallisation, the remaining cholesterol was converted to the 3-acetate and the specific activities were adjusted to the equivalent weight of cholesterol.

reviews

Economy and ecology

Peter D. Moore

The Economy of Nature: A Textbook in Basic Ecology. By Robert E. Ricklefs. Pp. v+455. (Chiron: Portland, Oregon, January 1976.) $13.95. *Methods in Plant Ecology.* Edited by S. B. Chapman. Pp. viii+536. (Blackwell Scientific: Oxford and London; Halsted: New York, 1976.) Cloth £15; paper £9.75.

PUBLISHERS of ecological textbooks would do well to consider Gause's principle of competitive exclusion. The cash in a student's pocket is a distinctly limited resource for which textbooks must compete; any new text, to be successful, must carve out a niche for itself in what is already a fairly saturated market. And in so doing it will inevitably displace other texts from that volume of literary hyperspace which they currently occupy. When considering a new text, one must therefore assess its competitive ability in the niche for which it has been designed.

Ricklef's new book is aimed at bridging the gap between the voluminous tests which seek to cover the whole gamut of ecological thought and the brief, sometimes superficial, introductory books. It has a broad coverage, following a very similar sequence to his previous book (*Ecology*, Thomas Nelson, 1973), but the text penetrates far less deeply than that volume. The style is light and readable and is uninterrupted by references to original papers. Some new illustrations, both verbal and pictorial, are introduced, although many of the old ones are retained. There is still a strong emphasis upon zoological and North American examples of the principles described.

Is there a niche for this textbook in the current market? I do not consider it a serious competitor with the variety of good textbooks currently available (including Ricklefs' *Ecology* and also Colinvaux, *Introduction to Ecology*, Wiley, 1973; *Ecology*, Krebs, Harper and Row, 1972; *General Ecology*, McNaughton and Wolf; Holt, Rinehart and Winston, 1973) for that portion of the 'resource' provided by British University students of biology. The low level of numerical treatment and the lack of references will prob-

ably cause it to fail in this respect. It is not a textbook which generates a critical approach to the data and the conclusions presented; therefore, I do not think that it will displace any weightier text from this portion of its niche. It may appeal to the broader-based groups of students possibly proceeding along paths encompassing humanities and environmental studies. Here its neglect of the numerical approach may be an asset. In this narrower niche, however, is the resource adequate to support a large population—that is how big is this specialised market? Where competition is weak it is often, although not invariably, due to low resource availability.

The second book, edited by Chapman, is making an overt bid to fill a specialised, relatively narrow niche in the textbook community, that of methodology in plant ecology. A wide range of disciplines, and methods within these disciplines is covered and, since each has a separate author, one might expect a degree of patchiness. The chapter on history of vegetation covers an enormous range of ideas, methods and techniques, but never in sufficient depth to provide an adequate laboratory manual. The key to pollen grains is inadequate, inaccurate and misleading. The description of vegetation is dealt with far too briefly, in the following chapter. Several good textbooks are now available on this subject and this short survey adds little new.

The chapter on production ecology is useful in that it explains very clearly and precisely the basis of productivity estimation and the various growth formulae that can be used. Once again it provides background information and methodological appraisal, but space (presumably) does not allow the detailed descriptions of techniques which would be required by neophytes.

Other topics covered include physiological ecology, soils, climatology, chemical analysis and data collection systems. In almost all cases the sheer pressure of space results in the abbreviation of technique description, although references to the literature are always numerous. The imposed brevity, however, does mean that the serious student of any specialised discipline is likely to look elsewhere for information, often to a text specifically on the required techniques. Competitive displacement is therefore an unlikely outcome of this publication. There remains the question of whether the gathering of all these varied techniques under a single cover might provide the book with a distinct niche of its own. I fear that this is unlikely; it may find its way into libraries or on to a few private shelves, but I cannot envisage its being bought by students; and this is where the real resource lies. ☐

Peter Moore is a lecturer in the Department of Plant Science at King's College, University of London, UK.

Pollution control

The Estimation of Pollution Damage. (Studies in Environmental Pollution.) By P. J. W. Saunders. Pp. 126. (Manchester University: Manchester, April 1976.) £6.95.

NO-ONE approves of pollution. Unfortunately, however, its control is not simple. A great deal of effort and money may be devoted to reducing the amount of some toxic substance in our environment, with little or no useful effect, because preliminary studies have not been made to find out what, if any, harmful effects the

substance is really producing. Our resources are clearly insufficent to deal with every possible source of pollution, and it is therefore essential that we discover which are the greatest danger, and concentrate, first, on their control.

In this book Dr P. J. W. Saunders of the Natural Environment Research Council, and formerly of the Pollution Research Unit at Manchester University, sets out to show how the real, harmful effects of pollution may be measured, so that effective control measures may be planned. The book

is primarily intended for students and administrators concerned with this subject, although it can also be recommended to the more studious members of the wider public who are often so vocal about environmental problems.

Dr Saunders defines pollution as "the introduction by man of waste matter or surplus energy into the environment which directly or indirectly causes damage to man and his environment, *other than* himself, his household, those in his employment or those with whom he has a direct trading relationship". I am unable to understand why self-inflicted, domestic and occupational exposures are excluded except where they interact with other types of pollution. It may be more difficult to evaluate this type of damage in cash terms but they

are equally serious and equally necessary to control.

The amount of information condensed into this small book is quite remarkable. The whole field of air, water, soil and of estuaries, and the specific problems caused by persistant pollutants is summarised in under 20 pages. Global problems are dealt with in about the same space. The remainder of the text describes the many techniques involved in estimating and measuring pollution, with a number of well-chosen practical examples. The extreme compression does not make for easy reading, which may discourage less-informed readers. Students who often only receive a superficial introduction to this subject will benefit most from its content.

Kenneth Mellanby

Changing spectrum for crop plants

Evolution of Crop Plants. Edited by N. W. Simmonds. Pp. xii+339. (Longman: London and New York, 1976.) £14.00.

THIS is an excellent book in which the title is used in its widest sense to convey the whole spectrum of a crop's development from origin to modern varieties, and its present and future potential role in the changing demands of world agriculture. Not since De Candolle has a book attempted to cover all the major and many minor crop plants grown in the world today; this book should therefore prove invaluable to all those involved internationally in crop improvement programmes.

The book is multi-authored, with 86 chapters detailing individual crops, arranged alphabetically by family, with the last two covering timber trees and minor crops. Each has an introduction and four sections on cytotaxonomic background, early and recent history and prospects, ending with a bibliography listing key and source references. Chapters range from 2,000–6,000 words depending on agricultural importance and depth of evolutionary understanding.

The absolute merit of individual chapters must be judged by the relevant crop specialist. All combine a range of scientific disciplines in framing a crop's history in a concise and comprehensive manner. In some instances emphasis on the former has produced some formidable terminology in the cytotaxonomic section (for example, Papaya, ch 8, and Fig, ch 59) and early history section (in

describing early plant migrations); and simplification or additional explanation would be helpful, although in both the regular use of diagrams aids presentation. The recent history and prospects sections detail the application of modern breeding methods in variety development, and attempt to identify the inherent problems for the future. In some cases, however, little mention is made of the local importance of a crop outside its main production area, where different though equally important problems may occur—two examples being cotton and groundnuts in Africa (south of the Sahara).

Of particular note for the future is the recurring theme in many chapters, on crops with a long breeding history, of the poor management in the past of crop genetic resources, resulting in depletion of available genetic variability, which may limit future advances. That this should not be repeated is stressed by a number of authors describing crops which have recently received attention and which, although only now grown in subsistence agriculture, are likely to play a key role in the future in improving living standards in developing countries. It is also evident that there are crops, such as cucurbits and edible aroids, whose potential for production in marginal tropical areas is recognised but as yet little exploited.

I can only agree with the editor's words that the book "concentrates on the particular and, in doing so reveals . . . , how insecure our knowledge often is, how much more work is needed and how often, even now, the right questions have not yet been asked".

H. E. Gridley

H. E. Gridley is a research assistant in the department of applied biology at the University of Cambridge, UK.

Foraminifera

Recent Foraminifera. By Esteban Boltovskoy and Ramil Wright. 515 pp. (Junk, The Hague, 1976.) 125 DFl.

A THOROUGH revision and updating of Boltovskoy's original *Los Foraminiferos Recientes* (Buenos Aires, 1965) has resulted in a comprehensive new text, furnished with well chosen and clearly reproduced figures, a carefully compiled bibliography and a good index, the whole being attractively printed on good paper and sturdily bound. The authors deserve high praise for so successfully summarising such a wide range of information, for always emphasising the significance and practical value of each aspect of the subject, and in producing such a readable book. It can be read from cover to cover with pleasure, or simply dipped into with profit.

No previous knowledge of the foraminifera is assumed, yet the authors quickly introduce the reader to the results of research which were very new when their manuscript was completed (January, 1975). Foraminiferal morphology and biology are clearly and concisely described, the scientific significance of foraminiferal studies is made plain, and the distribution and ecology of the foraminifera are described and discussed in their oceanographic, biostratigraphic and palaeo–oceanographic contexts. Taxonomy, systematics and problems of synonymy are dealt with concisely and realistically. There are most useful chapters covering techniques of collection, preparation, culture and storage, which the authors have tried and tested.

This book is no mere library compilation of other scientists' publications: rather, it is the result of long experience of practical work with foraminifera by the authors themselves. The book is intended for workers of all kinds, from the beginning student to the academic and industrial researcher, for those who have access to large research budgets and to those who can lay hands on only the most simple equipment and most modest resources. The authors succeed admirably.

Of course, there are the inevitable minor errors (a dreadfully confused Fig. 117 is, perhaps, the worst). The book should, however, be compulsory reading for all students of the foraminifera, be they protozoologists or micropalaeontologists, biological oceanographers or biogeographers.

F. T. Banner

Dr Banner is a reader in the Department of Oceanography at University College, Swansea, UK.

Understanding invertebrates

A Functional Anatomy of Invertebrates. By V. Fretter and A. Graham. Pp. vi+589. (Academic: London and New York, March 1976.) £12.50; $31.

THIS is a book meant to be read. It is possible to do so from one end to the other by those who know something about invertebrates; and the experience is refreshing. But when I finished it I wanted to write a different book about invertebrates. This somewhat ungracious comment is a compliment in a way, for those of us who know the authors' many contributions to our understanding of invertebrates will be neither surprised nor disappointed by this book.

It is not possible in so short a space to argue about what is meant by 'functional anatomy' for rather like the words 'physiology' and 'biology', it is interpreted in different ways by different people. So this is a peculiarly individual book and we must be grateful to them for so elegantly setting down their approach after a life-time's experience of teaching and research, and to Academic Press for producing it so well.

How fortunate students are; with Barnes, Meglitsch, and Marshall and Williams as basic texts, and with Barrington, and now this new book, it is possible to obtain a far better understanding of the invertebrates than ever before. I must emphasise the inherent readability of this new book, for so many basic textbooks are so packed with detail that they serve more for reference, or at least can be read with pleasure only in parts.

Fretter and Graham have given us an account of all the more important invertebrates (but excluding protochordates); although exemplary in style and clarity, the illustrations are not numerous and serve to illustrate the text. This is a Koran for the converted and in my opinion it will be appreciated most by those whose interest has already been caught.

How wise it was of Sinauer to put the stereoscan picture of *Didinium* attacking *Paramoecium* on the cover of E. O. Wilson's *Life on Earth.* How much more exciting it would have been in the present case to have ultrastructural details of the coelenterate nematocyst for example, or a stereoscan view of the choanocyte. What a pity there is so little on parasitic protozoa with no life cycle diagrams. Only seventeen pages on land arthropods. No illustration (even) of the cephalopod eye or the octopus brain. I must confess that if I had just left

the sixth-form having used Roberts, Berrill, and Clegg I would think this book, on first impression, was dull, but I would appreciate it more I think, when I had learnt more about the invertebrates. For there is no 'ideal' book to cover so wide a field; there can be no single volume which can provide all a student needs, especially when most of us devote less of the degree course to such basic zoology.

Published at a price most students will be reluctant to pay, it is nevertheless a book no Department can afford to be without. It is a book which provides a stimulating comparison with Barrington and others, and which, I would predict, will be appreciated more and more by students as they come to know the invertebrates better.

R. P. Dales

R. P. Dales is Professor of Zoology at Bedford College, University of London, UK.

Caribbean–Mexican tectonics

The Ocean Basins and Margins. Vol. 3: The Gulf of Mexico and the Caribbean. Edited by A. E. M. Nairn and F. G. Stehli. (Plenum: New York, 1975.)

THIS is the third in a series intended to document the world's ocean basins and their margins. Previous volumes have covered the North and South Atlantic Oceans.

The Gulf of Mexico and the Caribbean Sea have been the subject of many geological and geophysical studies that have added to the wealth of data obtained by the oil industry in the shelf and onshore areas. Not surprisingly, however, there is no firm concensus on the tectonic evolution of the area, although two views are held. The mobilist view seeks to explain the evolution of the Caribbean by seafloor spreading and plate tectonics, whereas the stabilists believe that the Caribbean Sea is an ancient oceanic structure.

There are fifteen chapters in the book. The introductory chapter by Uchupi reviews the Physiography of the Gulf of Mexico and Caribbean Sea. The two succeeding and particularly good chapters discuss Geophysical Studies in the Gulf of Mexico (Martin and Case) and Geophysical Studies in the Caribbean Sea. Fox and Heezen present valuable rock dredge data, and seismic refraction and reflection studies in the Geology of the Caribbean Crust. Other chapters discuss various aspects of Palaeozoic geology around the Gulf of Mexico and Caribbean Sea. There

are detailed reviews of key subjects such as the Northern Termination of the Andes and the Palaeozoic and Mesozoic tectonic belts of Mexico and Central America. Areas discussed in more detail are Cuba, Hispaniola, Jamaica and the Nicaraguan Rise. The Lesser Antilles and Aves Ridge are reviewed by J. F. Tomblin with emphasis on the volcanology, petrology and geochemistry of the arc. T. W. Donnelly concludes the book with a review of critical problems in understanding the Gulf of Mexico and Caribbean Sea.

Like the preceding volumes, this book has its deficiencies. Some of the chapters are little more than dry structural and stratigraphic descriptions of particular areas with little comment on the wider implications. Time has invalidated earlier conclusions—in one case beds previously considered as Permian are now known to be Late Cretaceous. In spite of this criticism, the chapters will be valuable as source material.

A more serious criticism of the book is the absence of any single detailed discussion of the marine geology of the Barbados Ridge and Lesser Antilles arc. Reportedly, wells on Barbados show evidence of imbrication, of obvious relevance in subduction. The Deep-sea Drilling Project results are cursorily treated and would have merited a whole chapter.

One of the authors rightly comments that a synthesis of the land and marine data is required and another that there is a "negative correlation between Caribbean field geology experience and readiness to apply the recent concepts of seafloor spreading". A useful example is the problem posed by the Jurassic igneous rocks of La Desirade situated in the Tertiary Lesser Antilles Arc. Are these rocks *in situ* or are they a tectonically transported slice of old Atlantic ocean crust? In such areas multichannel seismic surveys may resolve such tectonic anomalies.

Multichannel seismic surveys emerge from the book as priority measures for the Gulf of Mexico and Caribbean Sea to resolve such questions as continuity of salt across the Gulf and the distribution of beds below horizon B'' in the Caribbean itself. Such surveys, calibrated by deep-sea drilling, may resolve many of the problems of the stabilism-mobilism controversy, reviewed and briefly discussed by the contributors to the book. **David G. Roberts**

David Roberts is a Principal Scientific Officer at the Institute of Oceanographic Sciences, Godalming, UK. He was recently Co-Chief Scientist on Leg 48 of Glomar Challenger at the Margins of Biscay and Rockall as part of the Deep-Sea Drilling Project.

announcements

Appointment

Mr E. S. Booth, CBE, MEng, FRS, Chairman of the Yorkshire Electricity Board, Leeds, will take over as President of the Institution of Electrical Engineers for 1976/77.

Awards

The Association for Radiation Research has awarded the 1976 Weiss Medal to **Dr Ulrich Hagen** of the Institute of Radiobiology in the Nuclear Research Centre, Karlsruhe for his work on the effects of radiation on biological molecules.

The University of Chicago has awarded the first George Willard Wheland award to **Professor M. J. S. Dewar** of the University of Texas for his achievement in introducing new concepts from quantum mechanics into university courses on organic chemistry.

Meetings

November 8–10, **Function and Biosynthesis of Lipids** (Sponsored by the IUB and PAABS), Sierra de la Ventana, Argentina (Dr Nicolas G. Bazan, Chiclana 508,8000 Bahia Bianca, Argentina).

November 11–12, **Scientific and Technical Publishing in a Multilingual Society,** Luxembourg (Mr J. M. Gibb, DG XIII, European Centre, Kirchberg, Luxembourg).

November 22–26, **Corrosion Problems, Prevention and Practice,** Auckland, New Zealand (Conference Chairman, Australasian Corrosion Association, PO Box 5961, Wellesley Street, Auckland, New Zealand).

December 10, **New Aspects of Epithelial Transport,** London (The Secretary, Dept of Physiology, King's College, London).

December 28–31, **Coconut Research and Development,** Kasaragod, Kerala State, India (N. M. Nayar, CPCRI Regional Station, Vittal 574 243, Karnataka State, India).

January 11, 1977, **Corrosion in Agriculture,** Nottingham (B. Wilton, School of Agriculture, Sutton Bonington, Loughborough, Leics.).

January 26–28, 1977, **Electrical Properties of Biological Polymers, Water and Membranes,** New York (The Conference Director, The New York Academy of Sciences, 2 East 63rd

Person to Person

Myocrisin (sodium aurothiomalate) is said to photosensitise the skin of some patients. Any information concerning such an effect in humans or animals treated with drugs containing gold would be appreciated by Dr R. J. Wilkins, Dept. of Microbiology, University of Otago Medical School, Dunedin, New Zealand.

Smithsonian Fellowships will be awarded to support independent research in residence at the Smithsonian Institute using the collections, facilities, and laboratories and pertaining to research interests of the Smithsonian research staff (anthropology, biology, the earth sciences and the history of science and technology). Fellowships worth $10,000 p.a. and research allowances, are granted to postdoctoral scholars, and predoctoral fellowships, of $5,000 and research allowances, may be granted to doctoral candidates. Applications are due by January 15, 1977.

A limited number of Research Fund grants, (of ~£100) will be awarded for 1977. Application forms, together with the regulations governing the awards, may be obtained from the Education Officer, The Chemical Society, Burlington House, London W1V 0BN, and the closing date for applications is November 1, 1976.

Three awards (consisting of a silver medal and 250 guineas) for work in chemistry will be made to the British chemists who have published the most meritorious contributions in their fields of experimental chemistry. The rules for the award may be obtained from the Local Activities Officer of the Chemical Society, Burlington House, London W1V 0BN. Applications or recommendations must be received by December 31, 1976.

There will be no charge for this service. Send items (not more than 60 words) to Martin Goldman at the London office. The section will include exchanges of accommodation, personal announcements and scientific queries. We reserve the right to decline material submitted. No commercial transactions.

Street, New York, New York 10021).

April 18–20, 1977, **Chemical and Molecular Lasers,** St Louis, Missouri (Deadline for abstracts: January 10, 1977) (Dr W. Q. Jeffers, Helios Inc., PO Box 2190, Boulder, Colorada 80302).

April 18–22, 1977, **Molecular Beams,** Amsterdam (A. E. de Vries, FOM-Institute, Kruislaan 407, Amsterdam, The Netherlands).

April 20–27, 1977, **First International Conference of Science Editors,** Jerusalem (Miriam Balaban, PO Box 4059, Jerusalem, Israel).

May 4–6, 1977, **Ozone Technology,** Paris (Deadline for abstracts: October 31) (International Ozone Institute, European Committee Secretariat, 52 rue d'Anjou, 75384 Cedex 08 Paris, France).

July 24–30, 1977, **40th Annual Meeting of the Meteoritical Society,** Cambridge, UK (Dr E. R. D. Scott, Department of Mineralogy and Petrology, Downing Place, Cambridge CB2 3EW).

August 15–19, 1977, **International Conference on Applied General Systems Research: Recent Developments and Trends,** Binghamton, New York (George J. Klir, School of Advanced Technology, State University of New York, Binghamton, New York 13901).

August 30–September 1, 1977, **Dopamine** (ISN satellite symposium) (Dr P. J. Roberts, Department of Physiology and Biochemistry, University of Southampton, Southampton SO9 3TU, UK).

September 13–16, 1977, **Unconventional Photographic Systems,** Oxford (G. W. Smith, Ozalid (UK) Ltd, Langston Road, Loughton, Essex, UK).

September 14–16, 1977, **Quantum Electronics,** Southampton (Deadline for papers: May 31) (Dr H. C. Hanna, Department of Electronics, University of Southampton, Southampton, Hants SO9 5NH).

September 28–30, 1977, **Platelets: a Multidisciplinary Approach,** Florence (Dr G. de Gaetano, Laboratory for Haemostasis and Thrombosis, Instituto di Richerche Farmacologiche 'Mario Negri', Via Eritrea, 62–20157 Milano, Italy).

October 7–11, 1977, **Essential Oils,** Kyoto, Japan (Mr Yasumasa Kato, Secretary General, VII International Congress of Essential Oils, c/o Kyoto International Conference Hall, Takaraike Sakyo-ku Kyoto, 606 Japan).

Too little guidance on the PhD?

TWO-AND-A-HALF years ago the Expenditure Committee of the House of Commons put out a report on postgraduate education in Britain. Many of the recommendations stirred up strong feelings in the academic community, and there was considerable interest in how the Department of Education and Science (DES) would field this full frontal assault on the nature of the PhD. The recently published White Paper 'Postgraduate Education' (HMSO; 28p) is the rather belated reply.

The Expenditure Committee's most headline-catching suggestions were in the field of postgraduate financing. It had suggested that there should be at least an element of repayable loan in student awards above some basic minimum. This idea finds little favour with the DES and there seems no reason to resurrect it in the near future. But it was not in student finances that the strongest feelings were aroused; it was in the proposals that "pre-experience" students should be discouraged from taking non-vocational postgraduate courses but encouraged into vocational courses leading to qualifications which would be a prerequisite for entry to a career. These suggestions stemmed from the committee's view that "postgraduate education should be shaped, not by student demand alone, but principally by the needs of the economy and of society as a whole".

The Government "broadly accepts" this view in its reply; resources for postgraduate education should be provided "primarily in order to meet the country's need for trained manpower". It accordingly defends itself by pointing to the many moves that are being made to shape postgraduate education towards vocational ends, and remarks that "it would be an over-simplification to regard research training leading to the PhD as non-vocational". And yet having apparently gone a long way down the road to manpower planning in the postgraduate sector, the government then declares that "the content and balance of postgraduate training, in all its complexity, are best left to the autonomous interaction of the institutions themselves, the research councils and employers". It looks suspiciously inconsistent with manpower planning to expect that universities can work out individually how to do their bit to meet the country's and society's need for trained manpower; and it suggests that the DES, for all its waiting to hear the views of (unnamed) relevant bodies before venturing to put its own position forward, has succeeded only in compiling a hurriedly cobbled-together case for seeming to worry about the needs of the nation while actually defending the *status quo*.

In the face of the committee's assault, the DES should have come out fighting. What evidence is there that manpower planning makes sense when it comes to controlling the intake of prospective PhD students? The Expenditure Committee found a few representatives of industry prepared to say that PhDs didn't interest them; but if they had looked in other directions they would have found industrialists prepared to say just the opposite. The Civil Service, another major employer, would equally not be of one mind. With such variability in opinions there seems little sense in trying to evolve any major strategy to switch the emphasis in postgraduate work. It is difficult to believe that there is anything like enough evidence to allow the DES to start matching postgraduate education to the "needs of the economy and of society as a whole".

A somewhat worrying omission from the DES's response is any obvious rejoinder to the committee's view that the number of pre-experience students should be "considerably" reduced. It is indeed true that the number of awards being granted for postgraduate research is slightly in decline at present, and one wonders what it will lead to. Will the downward trend be allowed to continue even if and when the economy picks up? The DES paper mentions only "some areas of study such as mathematics or non-clinical medical research" as being hindered by a student having to take a break after a first degree. It then goes on to promise a higher proportion of post-experience students. The whole looks a little ominous. Maybe we have too many young PhD students; maybe manpower planning would work. One would need more than this White Paper to be convinced, however. The DES, in its stern effort to show what it is doing for the GNP, might too easily abandon education for training.

Flashes in ashes

Allan Piper considers the biggest and most promising of five new energy projects sponsored by the International Energy Agency (IEA).

THE opening earlier this week of a London headquarters for a new National Coal Board (NCB) subsidiary will probably cause no more than a passing ripple in the whirlpool of international energy affairs. Few people can have heard of NCB (IEA Services), and fewer still will know what it represents. But its potential long-term impact on the global energy scene should not be underrated.

In the rush to develop new energy technologies, coal, once in decline as a major energy source, has assumed a renewed importance. Under its Ten Year Plan the NCB aims to boost UK coal production by 15% over the next decade. The US plans to expand its coal industry by a factor of three. Elsewhere—in Europe, the Soviet Union and Australia—similar production drives are getting under way.

But bringing coal out of the ground is one thing; how best to use it is another. The value of coal lies not only in its potential as a fuel. It also shows massive promise as a chemical feedstock, and as a source of premium grade hydrocarbons. Much research is needed—and it is here that NCB (IEA Services) comes in. Set up by the NCB last Novermber, it will look after coal research in Britain for the IEA. Five projects, costing an overall £15 million, have already come its way. Four of them are relatively modest: costing an annual £1 million or so, they involve merely the collection of information for the next five years, keeping the ten nations involved abreast of economic and technological developments.

The fifth project is by far the largest and most promising. It involves the design, construction and operation of an 85 MW fluidised bed combustor on an old NCB site at Grimethorpe in Yorkshire. Jointly financed by Britain, the USA and West Germany, the combustor is expected to cost £6 million to build, and a little over £500,000 a year to run during its eight-year life as an IEA test bed—a total of around £10.5 million.

With so much attention being focused in Britain on nuclear power and North Sea oil resources, the true significance of Grimethorpe is likely to pass unnoticed. In the past, fluidised combustion—an elegant technique for burning fuels efficiently, economically and cleanly—has been sadly undersupported in relation to its technological potential. But Grimethorpe marks a turning point. Through the Depart-

ment of Energy (DEN), which will provide 66% of the cost of Britain's share of the project (the rest will come from the NCB), the UK Government is providing solid backing for the first time.

There are, moreover, preliminary discussions under way between the NCB and leading industrial companies on the early introduction of a 60 MW commercial prototype in addition to the Grimethorpe test bed. Such a step should be technologically possible, and might have been taken already had funds been available. As it is, success in the present talks, and subsequent government approval, could see the first British prototype operating within five years.

Bed of ashes

Fluidised combustors boast advantages that cry out for commercial exploitation. Because they are more compact and efficient than conventional burners, capital construction costs and running costs are lower. They operate at lower temperatures, forming less clinker and slag and thus lowering maintainance costs. They offer operational flexibility: multifuelled combustors can be designed to burn coal, oil and gas, with switchover times measured only in hours. Thus, while they might usefully exploit fluctuations in different fuel prices, they can also remain immune to temporary shortages of particular fuel types. Moreover, by burning extremely low grade coals they could help stretch economic reserves, reducing overall coal production costs.

These benefits spring from the way fluidised combustion works. The underlying principle is startlingly simple. In a fluidised combustor the proportional amount of air involved in fuel burning far exceeds that found in conventional burners. During combustion, tiny fuel particles are continually fed into a bed of fine mineral ash, maintaining a proportion of 1 part fuel to around 200 parts ash. At the same time, an even flow of air is passed upwards through the fuel bed, causing it to act like a turbulent fluid—hence the name fluidised combustion.

As the fuel particles mill around in the airstream they do not stick together because of the low combustion temperature. Consequently they burn extremely thoroughly, distributing heat more efficiently than under ordinary burning conditions. The heat can be utilised either by passing cooling tubes

through the fuel bed, or by using more air than is needed to maintain turbulence, and extracting it from the top of the combustion chamber.

The potential range of application is enormous. As electrical power generators, fluidised combustors might usefully be considered by the Central Electricity Generating Board (CEGB), now busily closing dozens of small, urban coal- and oil-fired plants that could perhaps be converted to complement the nuclear baseload. In industry, fluidised combustors can be used to run small boilers and drying kilns, and pressurised plant can drive gas turbines. They can also be designed as incinerators for sewage sludge, organic and chemical wastes, and everyday municipal rubbish.

With every application the waste heat is easily harnessed. District heating is an obvious area for development here. The DEN, now devoting greater attention to the potential of combined heat and power schemes, is unlikely to overlook this bonus.

The environmental advantages will not be overlooked either. In fluidised combustion the production of SO_2 and nitrogen oxides is easily controlled and minimised. This is because burning temperatures are in the relatively low range of 750–950 °C. On the one hand, this is too cool for the production of most oxides of nitrogen; on the other, it is perfect for the harmless combination of SO_2 with ground limestone placed in the ash bed.

Dusty record

Against that background the poor record of support for research and development (R&D) seems surprising. The CEGB, with potentially so much to gain, withdrew almost entirely from the
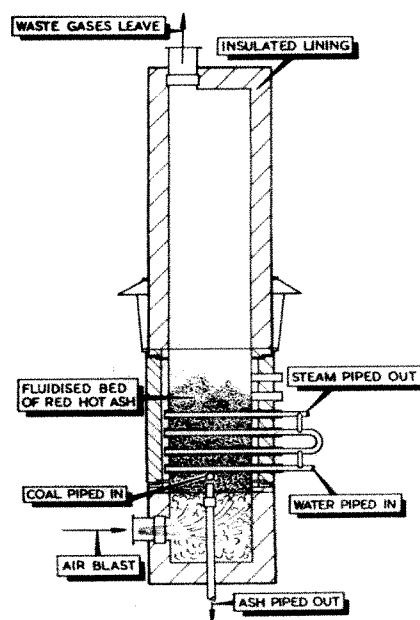
Fluidised bed combustor

scene soon after launching pioneering work in the middle 1950s, turning instead to the greater promise then offered by nuclear power and oil. It has since remained involved only in an advisory capacity. Throughout the 1960s the NCB carried the R&D effort virtually alone. The UK Government did not become directly involved until about five years ago, and then only in a minor way with the formation of Combustion Systems Limited (CSL), a tripartite research organisation involving the National Research Development Corporation, the NCB and British Petroleum (BP). The private sector has since shown increasing interest.

The Grimethorpe test bed is not the first in Britain. Progress made by the NCB during the 1960s led to the construction and successful operation of several small experimental rigs at the British Coal Utilisation Research Association (BCURA) in Surrey, and the Coal Research Establishment (CRE) in Gloucestershire. The BP Research Centre at Sunbury on Thames has also been the centre of recent applied research by CSL. Two small commercial plants are already operating in Britain, one of them at Renfrew in Scotland, where the enginering company Babcock and Wilcox runs a 13.5 MW fluidised combustion boiler. The other is used as a grass dryer by an enterprising Lancashire farming family.

Grimethorpe will not represent the first attempt at international collaboration on fluidisation combustion. Work at BCURA in 1970 and 1971 on environmental benefits was sponsored jointly by the NCB and the US Environmental Protection Agency. Over the past five years, following the withdrawal of NCB funding for BCURA, studies on pressurised plant have been supported entirely by the US Office of Coal Research. CSL facilities are currently used by the US Electrical Power Research Institute and the National Swedish Environment Protection Agency, together with many independent commercial organisations. CSL projects are themselves partially supported by grants from the European Coal and Steel Community. With more effort at both national and international level, development might have passed beyond its present point, but the funding was not available.

Perhaps the greatest single setback came in 1971 when the NCB, having decided to stop carrying the R&D burden on its own, failed to win government backing for a planned 20 MW demonstration steam boiler. Only £2 million was involved, and the Advisory Council on Research and Development (then under the Ministry of Power) had just reported favourably on fluidised combustion; initial design studies for the project had been sup-

ported by government grants totalling £200,000.
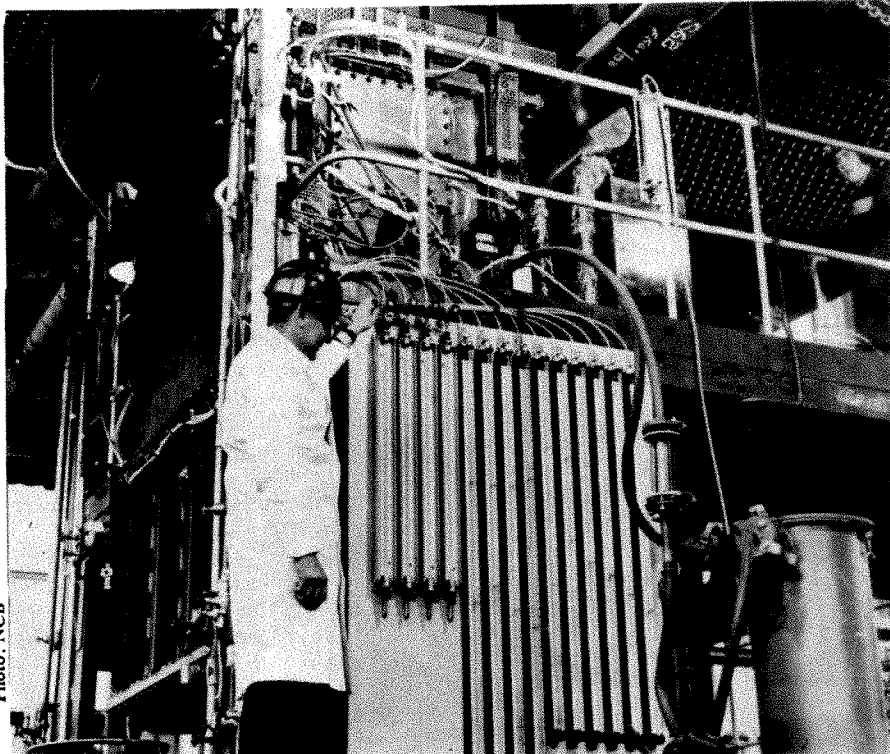
But the Tory government had only recently come to power, the CEGB, riding the tide of nuclear expansion, was indifferent about the idea and had itself declined to give support, and even the NCB commitment was less than total. Furthermore, export prospects did not look rosy: the US administration for the most part still spurned fluidised combustion, ironically preferring to place massive emphasis on the development of pollution control units for conventional burners.

Dead ahead

Following that setback it was arguably only through the formation of CSL that Britain maintained its technological lead. Not until the coal industry examination of 1974—involving the government, the NCB and mining unions—did fluidised combustion become identified as a priority area for R&D, along with pyrolosis and liquefaction. The decision to support Grimethorpe with DEN funds acknowledged that finding. Orders for hardware should go to UK, US or West German contractors within the next 12 months.

So fluidised combustion looks set to come into its own. In the UK, apart from Grimethorpe and the planned prototype, CSL is actively considering a 66 MW combined heat and power pressurised gas turbine. The British Steel Corporation may well introduce a small unit into its River Dunn works in Yorkshire. The General Electric Company has declared an interest in the

proposed British prototype. And a new joint company, Babcock–CSL has been established as a design and export base.

In the US too, R&D is pressing ahead independently of the IEA project. Work on a $15 million pressurised test bed is due to start near Chicago this autumn, sponsored by the US Energy Research and Development Administration. A lot is at stake. While coal makes up 90% of recoverable fossil fuel researves in the US, much of it has an extremely high sulphur content. Fluidised combustion offers the opportunity to exploit these reserves without violating anti-pollution standards.

The enthusiasm is matched in Europe, in spirit if not with financial backing. Although only West Germany will share the running costs of Grimethorpe with Britain and the US, the other seven member nations of the IEA Working Party on Coal Technology will have access to data. France, alone among the major European nations, has remained cool, but her absence poses no problems.

Fluidised combustion could have tremendous relevance to the future world energy scene. Few now question the crucial importance of nuclear power. And few will argue with the NCB's suggestion that coal's greatest potential lies in pyrolysis and liquefaction. But where coal must be used as a fuel, fluidised combustion looks like becoming the cheapest, cleanest, and most efficient way of using it. As one NCB board member has put it, the greatest innovations are yet to come. □

US ENERGY

Making oil out of money?

A controversial plan to pour billions of dollars of federal funds into commercial development of "synthetic" oil and gas from coal and from oil shale is about to face a crucial test in the US House of Representatives. Colin Norman reports

A SYNTHETIC fuels industry is a keystone of the Ford administration's long range energy research and development (R&D) strategy. The idea is to force-feed its growth in the United States to the point where it could begin to supplement dwindling domestic supplies of oil and gas in the 1980s and 1990s. The plan now arousing controversy is embodied in a bill already approved in differing versions by the Senate and by four separate House committees. It would provide up to $6,000 million to private industry over the next two years in loan guarantees and other financial assistance for the design and construction of large scale production plants.

According to the bill's supporters, chief of whom is Vice President Nelson Rockefeller, the measure is vitally needed to halt the growing imbalance between supply and demand from domestic oil and gas. Without the production of synthetic fuels, the argument runs, imports will inevitably increase as domestic supplies begin to peter out towards the end of the century. And without federal support now, there will be no synthetic fuels industry in place when it is needed.

To the plan's many opponents, however, the idea is a waste of money, an unnecessary subsidy to the already bloated oil industry, and a gross misapplication of energy R&D funds.

The latest five-year R&D plan of the Energy Research and Development Administration (ERDA), published last April, envisages a major role for synthetic fuels in the 1990s, suggesting that they "offer a domestic energy alternative to imported oil and gas". ERDA noted, however, that it will require production of between five and ten million barrels of synthetic fuel a day by 1995 simply to hold imports at their present level, and argued that "for the necessary number of plants to be operating in the mid-1990s, an industrial base on the order of 1 million barrels per day must have to exist by 1985".

ERDA's strategy is thus to press ahead with R&D on advanced technology for producing oil and gas from coal and shale, but at the same time it is urging industry to move rapidly ahead with the design and construction

of demonstration plants using technology already developed. The problem, however, is that such large scale plants are expected to cost about $1,000 million apiece, and, since the economics of the enterprise are difficult to predict, industry is having a tough time raising that sort of money. Hence the argument for Congress to provide loan guarantees and other financial incentives to help the industry get into the synthetic fuels business. ERDA officials describe the situation as similar to that faced by the nascent nuclear power industry in the 1950s.

Opponents of the bill have challenged its justification with arguments which go to the root of the Ford administration's energy policy. Led by Representative Richard Ottinger, a Democrat from New York, they support ERDA's programme for developing advanced synthetic fuel technology —ERDA is already funding a number of pilot scale projects jointly with industry—but they argue that it is premature for the Federal Government to pour vast sums of money into commercialisation efforts.

Those arguments received valuable support late last month from the General Accounting Office (GAO), an investigative body controlled by Congress. In unusually blunt language, the report recommended that "Government financial assistance for commercial development of synthetic fuels should not be provided at this time", and suggested that conservation efforts are likely to be much more cost effective.

The GAO study suggested that oil produced from coal or shale could cost between $18 and $30 per barrel, compared with the current price of about $12 for imported crude. It also estimated that synthetic gas would cost at least twice as much to produce as natural gas. Thus, to make the products competitive, the Federal Government would have to provide direct subsidies in the form of price supports or production assistance, the report suggests.

The GAO report was greeted with jubilation by the bill's opponents, coming as it did on the eve of the House debate. But ERDA officials were understandably less enthusiastic. For example, Robert C. Seamans Jr, the Administrator of ERDA, said in a statement that the report "presents strong conclusions and recommendations to the Congress without a sound underlying basis of analysis supporting them". He took issue with most of the report's conclusions, particularly its

comparison with future synthetic oil prices with present import prices, and suggested, finally, that its publication would be "damaging to the early implementation of our needed energy supply programs".

Equally disturbed was Representative Olin Teague, Chairman of the House Committee on Science and Technology and the chief architect of the bill. Teague called a committee meeting to discuss the report last week, and suggested that he would not bring the bill to the floor unless he is sure that he has sufficient votes to get it through. Later in the week, he said that he would try to get the bill before the full House as soon as possible, indicating that he feels it will pass.

As for the impact of the bill on the synthetic fuels programme, many officials see it as crucial. William T. McCormick, Chief of ERDA's Office of Commercialisation, said last week, for example, that "if this bill doesn't go through, many projects which have been supported by company funds will fold". He added that if the bill fails to pass this year, it will be very difficult to resurrect it next year.

Be that as it may, at least one synthetic fuels programme—the oil shale development effort—is already ailing so much that it may be difficult to revive it simply with a financial injection.

It should be recalled that not too long ago, the vast amounts of oil locked into shale beneath the Rocky Mountains in Colorado, Utah and Wyoming were being touted as the answer to America's energy problems. In January, 1974, the Department of the Interior began to offer small tracts of federal land to the oil industry for testing and development projects. The oil industry paid handsomely for the opportunity, the first 5,000 acre tract fetching a staggering $210 million.

Last month, however, consortia which hold rights to the tracts now leased in Colorado and Utah, withdrew from the programme and asked that the leases be suspended. And late last year, a consortium known as The Oil Shale Company (TOSCO), which was reckoned to be the leader in shale processing technology, announced that it had suspended its own private operations. They all cited the high capital cost of building the plants and a variety of environmental problems as reasons for pulling out.

If Congress passes the Synthetic Fuels Loan Guarantee Bill it may at least help the economics of shale processing, since it would reduce the cost of borrowing funds on the capital market. But, according to an Interior Department official, the prospects for reviving the effort look gloomy. □

GENETICS

Laying the guidelines bare

The long, tortuous process of setting controls on experiments involving recombinant DNA in the United States has taken yet another potentially important turn. Colin Norman reports

THE National Institutes of Health (NIH), the agency responsible for supporting most recombinant DNA experiments in the United States, last week published a 100-page document analysing the possible effects of the safety guidelines which it issued last June. Publication of the document, a legal requirement under the terms of the National Environmental Policy Act (NEPA), is likely to provide an important means of gathering public opinion on the guidelines, but it could also pave the way for some legal challenges to them.

Links between recombinant DNA research and environmental policy may seem a bit tenuous, but the language of NEPA is clear. It specifies that government actions which could have environmental implications should be preceded by an impact statement detailing the environmental risks and benefits associated with that action. Since microorganisms bearing recombinant DNA molecules could affect the health of plants, animals or man, the issuing of guidelines designed to keep such organisms in the laboratory and out of the general environment clearly comes under the terms of NEPA.

That fact was realised in the NIH while the guidelines were in the late stages of drafting, and caused considerable consternation in the agency, contributing to some delay in publication of the guidelines. In the end, however, Donald S. Fredrickson, Director of the NIH, decided to short-circuit the normal NEPA procedure by publishing the guidelines first and the environmental impact statement later. Therein lies a potential problem for NIH.

Already, NIH has received a number of complaints that Fredickson's action has circumvented public participation in the formulation of the guidelines. Among the complaints is a letter from two researchers at the Friends of the Earth, a national environmental organisation. The researchers, Francine Simring and Larna Saltzman, have asked NIH to stop funding all recombinant DNA experiments until the correct NEPA procedure has been followed, arguing that prior publication of the guidelines violates the spirit as well as the letter of NEPA.

Fredrickson has claimed, however, that prompt publication of the guidelines served the public interest better than following the usual NEPA procedure. Until the NIH guidelines were published on June 23, recombinant DNA experiments were proceeding under general guidelines drafted by an international group of geneticists which met at Asilomar, California, early in 1975. The NIH guidelines are in many respects more stringent than the Asilomar guidelines and they also establish a detailed implementation procedure. Thus, the impact statement itself suggests that "the escape of potentially hazardous organisms was more likely in the absence of NIH action". Though such arguments are rational enough, the possibility of a legal challenge should not be completely ruled out.

Nevertheless, the impact statement provides an important addition to the literature on the regulation of recombinant DNA experiments since it provides the first public discussion of the reasons why NIH chose not to follow some courses of action advocated by various groups and individuals. It also contains a discussion of the possible risks and benefits associated with the research, attempts to estimate the probability that some hazards will be realised, describes the guidelines, and outlines the procedure by which they were adopted.

As far as alternative courses of action are concerned, the statement discusses the following:

● No action—simply continue to rely on the Asilomar guidelines as the basis for regulating recombinant DNA experiments. Such a course was rejected because the Asilomar guidelines are considered too lax in some cases, and because there is no formal mechanism to ensure that they will be followed.

● Prohibition of funding from NIH for recombinant DNA experiments. That suggestion was rejected because other sources of funding are available, the potential benefits from the research would be foreclosed, it would "undermine American leadership in the establishment of worldwide safety standards", and the "leadership of the United States in biological research would be threatened". A flat prohibition would also interfere with free scientific inquiry, the statement suggests.

● Restrict all permissible experiments to a few maximum containment facili-

ties—a suggestion put forward most notably by Robert Sinsheimer, Chairman of the Department of Biology at California Institute of Technology. The statement notes that such a course of action would not distinguish between experiments with high potential risk and those believed to pose little, if any, hazard. It would also discourage many experiments and amount to a prohibition on some research because of limited access to such facilities, the statement suggests.

● Prohibit the use of the bacterium *E. coli* in recombinant DNA experiments. Because *E. coli* is a common inhabitant of the human gut, many observers have argued that it is a particularly risky choice as the host for transplanted genes. The statement points out that the particular strain of *E. coli* used in the laboratory has many features which limit its infectiousness. Moreover, it can be genetically altered to reduce its viability even further, as required by the guidelines for some classes of experiments. The statement also defends the decision to permit the use of *E. coli* by noting that no other bacterial species now seems to offer as many safety features.

● The statement also notes that some observers have questioned the decision to permit experiments involving the formation of recombinant molecules from uncharacterised pieces of DNA from warm-blooded animals—so-called shotgun experiments. It also notes that the use of DNA from oncogenic viruses has been challenged. The statement does not defend the decision to permit such experiments, however; instead, it merely notes the objections.

● The environmental impact statement has now been circulated in draft form to numerous scientists and other interested individuals and it is also available on request from the NIH division of environmental services. It is open for comments for 45 days, after which the NIH may convene a series of public hearings. Eventually, the document will be expanded and cast in final form. The comments will also be considered during any future revision of the guidelines.

Because both the risks and the benefits from recombinant DNA experiments are speculative, the impact statement is necessarily vague in places, and is also shot through with value judgments in areas where experimental data are lacking. Critics of the guideline and of the research will therefore find plenty to criticise in the document. A heated and lengthy debate should be anticipated. □

IN BRIEF

NRPB evidence published

The UK National Radiological Protection Board (NRPB), which advises the government on radiation hazards, has urged a fresh appraisal of Britain's approach to nuclear waste management. The board's evidence to the Royal Commission on Environmental Pollution (RCEP), disclosed last week, adds credence to the view that the RCEP's forthcoming report on radiological safety in Britain will itself criticise the relatively disappointing level of UK research into the safe disposal of long-lived radioactive nuclides arising from the nuclear power programme.

The NRPB evidence nonetheless questions recent attacks on international exposure limits for plutonium; and it endorses calls for individually established standards for radioactive discharge from nuclear installations, rather than blanket legislation.

AGR running smoothly

Barring unforeseen setbacks the commercial output of Britain's first Advanced Gas Cooled Reactor (AGR) nuclear power station, Hinkley Point B in Somerset, will have doubled to 1,000 MW by the coming winter. The second of Hinkley B's twin 660 MW reactors, now undergoing final egineering tests, is expected to come online in October, closely on schedule. The first has been performing smoothly at around 500 MW since July, following its switch on eight months ago—outstanding corrosion problems have restricted output to about 80% of the design level. Hinkley B is already producing electricity as cheaply as first generation, magnox nuclear installations, and could eventually become more competitive.

Though the Central Electricity Generating Board last week expressed renewed confidence in the AGR, which it regards as among the safest and 'cleanest' of existing nuclear reactor types, export prospects remain bleak because of its comparatively high capital cost and unproven commercial record.

Uranium fix?

Officials from California's State Energy Commission and Public Utilities Commission have in the past few weeks given the US Department of Justice and the Senate Foreign Relations Committee copies of documents obtained by the Australian Friends of the Earth (and published in the Australian press) apparently revealing the existence since 1971 of an international cartel to fix the price of free world uranium.

The documents, apparently obtained from the files of Mary Kathleen Uranium of Australia, are said to contain references to a meeting in Johannesburg in January 1974 attended by representatives from Australia, South Africa, Canada, France and the London-based Rio Tinto Zinc.

No official comment on the matter was available last week from the Justice Department in Washington, which in July served subpoenas on US and Canadian companies in connection with its investigations of a price ring, or from the Uranium Institute in London, which the documents reportedly suggest might have been intended as the successor to the Uranium Marketing Research Organisation, the name purportedly used by members of the "club" to describe their group.

Solar energy gathering

The planning and coordination of the international effort to develop solar energy was one of the subjects which about 150 scientists from more than 50 countries discussed last week when the World Meteorological Organisation (WMO) and UNESCO held a symposium in Geneva. The symposium, a sequel to one held in 1973 at UNESCO headquarters, also discussed problems, both technical and educational, associated with the conversion of solar radiation.

Nuclear exports overseer

Following the first meeting of the new French Government, the Elysée Palace last week announced the formation of a high-level government council to oversee France's participation in the growing international trade in nuclear technology. It will "define and coordinate sales of nuclear techniques and products".

President Valéry Giscard d'Estaing will head the new committee, which will also include the new Prime Minister, M Raymond Barre, the Foreign, Defence, Finance, Industry and Foreign Trade ministers, and the president of the French Atomic Energy Commission (CEA).

The creation of the new body comes after the call from non-aligned states at their recent conference in Colombo for an oil embargo against France. It also follows the row between France and the US over the sale of a reprocessing plant to Pakistan and the controversy over France's deal involving nuclear plants for South Africa.

Dr Henry Kissinger is meanwhile due to see President Giscard d'Estaing in Paris this week for talks which are expected to include discussions on French nuclear policy.

Biblis to restart

The 1,200 MW Biblis nuclear power station near Essen, West Germany, which was shut down in April after a routine check revealed cracks in the feed-water container, was expected to be back in operation this week, according to a spokesman for the operating utility RWE. News of the fault first emerged in early July, when RWE and Kraftwerk Union, which built the plant, disclosed the cause of the delay. Kraftwerk Union handed the twin pressurised water reactors over in February last year.

---

news and views

Harvesting whale and ·fish populations

from Robert M. May

FROM August 31 to September 9, the Advisory Committee on Marine Resources of the UN Food and Agriculture Organization is holding a Scientific Consultation on Marine Mammals in Bergen. The background papers for the meeting make a fair-sized pile. The major findings are summarised in the reports of four *ad hoc* groups, focusing on large cetaceans, on small cetaceans and sirenians, on seals and marine otters, and on ecological and general problems. A recurring theme is the difficulty of obtaining reliable estimates of population density, and of transferring estimates of population density in one area to another. For example, in reviewing population assessments of Antarctic fin whales, Gambel (conference document ACMRR/MM/EC/9) observes that the International Whaling Commission (IWC) has agreed on a current population estimate of about 83,000 for this species. Gambell notes that this estimate rests, *inter alia*, on a particular model, but that both the model and the numerical estimates of its key parameters are open to question. Moreover there appear to have been "sharp but erratic declines in the fin whale stock indices" in the areas designated III and IV by the IWC, and also a sharp decline in the catch-per-unit-effort in area V; these declines are not consistent with the IWC model. This erratic behaviour is typical of many exploited whale and fish populations, where the amplitude of population fluctuations increases as harvesting rates increase.

In theory and in practice, a central concept in the harvesting of a natural population is the maximum sustained yield, or MSY. The population may be thought of as having some intrinsic nett growth rate, d*N*/d*t*, which depends on the population density, *N*· at low densities, the per capita growth rate is relatively high, but the population is itself small, so that initially the overall growth rate increases as *N* increases; at high densities, the per capita growth rate typically falls, attaining zero when *N* reaches a value *K* (the "carrying capacity" of the environment), and

then becoming negative. Increases in per capita fecundity as population levels are lowered have been documented for several fish and marine mammal populations, although the underlying mechanisms are not always well understood. Possible explanations include the greater wealth of resources in a less crowded environment, and a variety of behavioural responses (some of which, in gregarious marine mammals, can stem from the breakdown in social organisation attendant upon the removal of dominant males or females). The upshot is that the typical plot of total population growth as a function of population density (d*N*/d*t* against *N*) rises from zero at *N*=0, attains some maximum value at *N*=*M*, and then declines back toward zero at *N*=*K*, thereafter remaining negative. A familiar equation which represents this phenomenon is the logistic d*N*/d*t*=r*N* (1−*N*/*K*), but this is only one of many equivalent forms (see for example Holt, ACMRR/MM/EC/29).

In its natural, unharvested state the population will be around the equilibrium value, where recruitment balances mortality, at *N*=*K*. Harvesting constitutes an additional source of mortality, and (provided the harvesting rate is not too high) the population will decrease to some new steady value at which the population gains from the intrinsic growth rate exactly balance the losses due to harvesting. The MSY is achieved by harvesting at a rate such that the population remains around *N*=*M*, that is, by harvesting at a rate equal to the maximum possible nett growth rate. The population may be thought of as a capital sum in a bank, with however the odd property that the interest rate depends on the magnitude of the sum. For sustained yield, one harvests only the interest on the capital; and for maximum sustained vield one arranges that the capital sum shall be that which generates the maximum interest rate.

Two papers in the collection call attention to shortcomings in overly simple notions of MSY, and make important extensions to the theory.

Holt, the convenor of this con-

ference, gives a catalogue of different formulae which have been used to describe recruitment curves. He shows that, on the one hand, available data are rarely of sufficient accuracy to discriminate among the alternative forms, but that, on the other hand, the different forms can lead to significantly different predictions as to MSY. In extreme cases, the harvesting rate that gives the MSY for one curve (for example, a curve skewed to low densities, such as the Beverton and Holt d*N*/d*t* = (1 − *b*)(1 − *N*)*N*/{*N* + (1 − *b*)*N*} with *b*=0 33) can result in extinction of the population if some other curve in fact pertains (for example a curve skewed to high densities, such as the Pella and Tomlinson d*N*/d*t*=b*N*(1−*N*) with *b*=0.67). Holt thus cautions against the uncritical use of MSY calculations to set quotas (see ACMRR/MM/SC/61, and also Allen, SC/57, and Gulland,SC/82). He recommends that a range of different quotas should be set for different stocks of the same species (in different geographical regions), these quotas being set on the basis of a spectrum of estimates of the state of the stock and the form of its recruitment curve; in this way we may hope to learn more.

Holt and the *ad hoc* group IV (ACMRR/MM/SC/5) also note the need to incorporate the effects of random environmental variability into simple population models. One outcome of preliminary explorations in this direction (Beddington, personal communication) is the suggestion that as harvesting rates increase, so too do the relative fluctuations in the population magnitude, and in the catch itself. This goes some way toward explaining an observed property of exploited populations, which was noted above.

Clark (conference document ACMRR / MM / SC / 65, and also *Science*, **181**, 630–634; 1973) makes the important observation that the aim of most commercial fish and whaling industries is not so much to maximise sustained yield, as to maximise the present value, PV, of (discounted) nett economic revenue. To this end, Clark takes the sort of considerations used

for MSY calculations, and modifies them in two ways. First, the harvesting costs are subtracted from the harvest, itself expressed in money units. Second, the present value of the nett economic revenue is obtained by integrating over all future yields, discounting at some appropriate discount rate, δ (which will usually be of the order of the bank interest rate, $\delta = 10$–20%). The aim is now to maximise PV. Although this approach is oversimplified in many ways (it neglects fixed capital costs for example), it is more realistic than classical MSY calculations in that it automatically incorporates the trade-off between short-term and long-term benefits.

Clark then shows that PV is maximised by a strategy which initially depletes the stock to some predetermined value, L, and thereafter harvests it for sustained yield (keeping $N = L$). His remarks are illustrated by numerical calculations for particular assumptions on recruitment and cost curves, but the conclusions are clearly general. Unless we have a combination of relatively low discount rate and relatively high harvesting cost, L will be less than the classical MSY point at $N = M$. The MSY approach will, however, give a good first approximation so long as the discount rate δ is significantly less than the intrinsic per capita rate of population growth, r, for the harvested species. Conversely, if δ is greater than, or of the order of, r then the sustained population value for maximum PV is significantly smaller than the MSY value; $L \ll M$. Clark observes that "international institutions such as the IWC seem to have been established on the assumption that the economic interests of the industry would, if properly channelled, automatically ensure the conservation of the resource". He emphasises that his "analysis indicates that the economic incentive for conservation of such resources may be quite minimal, as far as the commercial industry is concerned. The economic values of the whaling industry, of course, may not coincide with long term economic values of society as a whole".

If, for the moment, harvesting costs are neglected, then a simple but rigorous generalisation of Clarke's results can be given: maximisation of the present value of discounted nett economic revenue of a harvested population will permit the population to persist if, and only if, the maximum per capita population growth rate exceeds the discount rate. If not, PV is maximised by harvesting to extinction. In particular, if the overall population growth curve is logistic, the ratio of maximum PV to MSY sustained populations is $L/M = 1 - \delta/r$ if $\delta < r$, and zero if $\delta > r$. Returning to the

American Indians whaling in the 17th century

image of a fish population as a capital sum in a bank, we see that if the fish are not capable of an intrinsic per capita growth rate in excess of the bank interest rate, then we do better to turn the entire capital sum—the entire fish stock—into real money in a real bank.

For many commercial fisheries in the North Sea, r is indeed sufficiently large for conventional MSY calculations to retain validity as sensible first approximations. But for the Antarctic blue whale r has ben estimated at around 5%, which as Clark emphasises is less than (or at least well within) the normal range of discount rates. Consequently PV may be maximised by fishing such whale species virtually to extinction. The policies pursued by the whaling industries of countries like Japan are not the result of brutal stupidity, but rather are the optimal solutions of problems defined in narrowly economic terms.

Returning to the question of harvesting costs, we note that these will typically be higher per unit catch at lower population densities. Thus it is never economically rational to harvest the last few animals. There remains the worry that this lower economic breakpoint may be below some biological threshold population density, where for geographical or behavioural reasons the population is incapable of maintaining itself; ecologists refer to the existence of such a threshold as an "Allee effect". No such Allee effect has been demonstrated for any marine mammal, so that this economic consideration may contain one gleam of cheer.

Clark's work, which is developed further in his book *Mathematical Bioeconomics* (Wiley, New York, 1976),

is very illuminating. It undercuts the pious belief that the interests of conservation and industry can march together if only the world is seen clearly. The conservationist is essentially a person who does not discount the future, a person for whom $\delta = 0$; for the industrialist, δ is typically the prevailing bank interest rate. Their interests can thus be fundamentally irreconcilable, and no amount of methodological sophistication can help. The overriding decisions must be qualitative social and political ones, on the effective rate at which we discount the future of the world. Then, and only then, can we usefully employ the technical apparatus of MSY, optimum present value of discounted economic revenue, and the like. □

Polycyclic aromatic carcinogenesis

from S. Neidle

OF all the many external agents which can induce cancerous growth in animal cells, polycyclic aromatic hydrocarbons have been the most intensively studied, and indeed this topic has for long occupied a central position in cancer research. Just 200 years ago Pott first described the effects of soot and coal tar in producing a high incidence of scrotal cancer in chimney sweeps—it was not until the 1930s that Kennaway and Cook isolated in pure form the active chemical carcinogen from coal tar and identified it as benzo[a]pyrene, a polycyclic aromatic molecule with five condensed six-

membered rings, and an almost planar molecular geometry. Since then a large number of analogues both potent and inactive have been discovered and/or synthesised.

In spite of considerable research over the past two decades, the molecular basis of action of these compounds is still imperfectly understood, however. An initial problem still without a definitive answer concerns the cellular target of these carcinogens. Both proteins and nucleic acids have been implicated by various workers; however it is now generally agreed that DNA is probably the most likely candidate. Certainly, binding of carcinogenic polycyclics to DNA, albeit at exceedingly low levels, has been well established, largely through the extensive studies of Brookes, Sims and their collaborators. At first sight, it is attractive to conceive of this binding in terms of intercalation between DNA base pairs, as occurs with many planar drug molecules, especially since the hydrocarbons themselves are chemically unreactive. However, there is now clear evidence that the binding is covalent both *in vivo* and *in vitro*. Moreover, intercalation presupposes a planar molecular conformation for the substrate, and there are several examples of very active polycyclic carcinogens which are markedly non-planar and therefore could not possibly act as intercalators, as Glusker, Zacharias and Carrell have shown by X-ray crystallography in the case of 7-chloromethylbenz [a] anthracene (*Cancer Res.*, **36**, 2428; 1976). An alternative hypothesis by Boyland that binding involves an intermediate metabolite, such as an epoxide, has in recent years received widespread general acceptance, and it is now believed that the microsomal "mixed-function" oxidase system is responsible for this metabolic activation into a chemically reactive form.

A region of the polycyclic hydrocarbon molecule considered to be of likely accessibility to such activation is the so-called K-region. Indeed, Grover and Sims have shown that K-region epoxides do bind to DNA both *in vivo* and *in vitro* (*Adv. Cancer Res.*, **18**, 165; 1974), and are more active than the parent hydrocarbons. Weinstein, Harvey *et al* have investigated the *in vivo* binding of such an epoxide, 7,12-dimethylbenz [a] anthracene 5,6-oxide to various nucleic acids and synthetic homopolynucleotides (*Biochemistry*, **14**, 3451; 1975) and showed that there was a preference for reaction with the guanine bases in poly(G) They have now followed up this work with an elegant characterisation of the four major guanosine adducts obtained from separation of the hydrolysis products from the poly(G)–hydrocarbon complex (*Proc. natn. Acad. Sci. U.S.A.*, **73**,

Sounds volcanic

from Peter J. Smith

ACOUSTIC noise from volcanoes must have been heard for thousands of years; yet surprisingly perhaps, no recordings of volcanic sound appear to have been made before the 1950s (in Japan). The history of the scientific study of volcanic noise is not quite as short as that would imply, however. What were probably the first systematic observations were made at Vesuvius in 1906 by Perret, who later continued his work at Etna, Stromboli, Kilauea, Pelee and elsewhere (see *Volcanological Observations*, Carnegie Inst. Wash. Publ. 549, 1950). More recently, Wilcox (in *US Geol. Surv Bull. 965-D*, 1956) has described 11 distinct sounds from Paricutin, and Richards (*J. geophys. Res.*, **68**, 919; 1963), in a wider study, has correlated different sounds with the type of volcanic activity producing them.

But there is more to the study of volcanic sound than simply description and classification. Gorshkov (*Bull. Volcanol.*, **23**, 141; 1960), for example, used microbarograph records to estimate the energy discharged to the atmosphere by several large eruptions, the largest of which was the Krakatoa explosion of 1883 (8.6×10²³ erg). The Shiveluch eruption of 1965 was rather smaller, with only 1.8×10²¹ erg of energy emitted in air waves; but Gorshkov and Dubik (*Bull. Volcanol.*, **34**, 261; 1970) were able to take the analysis of this particular event even further. They found that only 0.10–0.15 per cent of the total kinetic energy derived from the eruption was imparted to the air. According to their estimate, the total energy of eruption was 1.3×10²⁴ erg, most of which was transmitted in the ejecta as thermal energy.

A second quantitative line of enquiry derives from the origin of the sound. The acoustic noise asso-

ciated with volcanic eruptions is, of course, due to the flow of gases—in particular to the interaction of gases with the stationary sides of the volcanic vent and to turbulence in the gases themselves. So as Woulff and McGetchin (*Geophys. J.*, **45**, 601; 1976) point out, analysis of volcanic sound should provide information about volcanic gas behaviour in general and about the velocity history of the gas in particular.

For example, a theoretical study carried out by Woulff and McGetchin suggests that three types of acoustic radiation from volcanic gas flow are possible—monopole, dipole and quadrupole. But the question then arises as to which of these occur in practice. Unfortunately, observational data are sparse; but analysis of recordings of the sound associated with fumarole activity at Volcan Acatenango, a stratovolcano in Guatemala, shows that in this particular instance the radiation is predominantly, if not entirely, dipolar. Extrapolation of this result to other types of volcanic activity is difficult, of course, although the little information available suggests that dipolar radiation should also be predominant in the more violent 'strombollian type of eruption, at least as long as the gas flow is subsonic.

Evidently quantitative studies of volcanic sound are still in their infancy. If acoustic data are to be translated into absolute gas flow velocities it will be necessary to know not only the type of radiation predominant in each particular case but also the constant of proportionality relating velocity to acoustic power in the power law appropriate to that radiation type. Again, the latter information is largely unavailable at present. As an intermediate step, however, there is the hope that even incomplete data will allow the elucidation of relative changes during an eruptive history and thus contribute to an understanding of eruption dynamics.

2311, 1976). In all four adducts, linkage with the polycyclic aromatic is by way of the 2-amino guanine group; two-diastereoisomers are linked to the 5-position of the hydrocarbon, and the other two, are similarly linked to the 6-position.

However, there is increasing and convincing evidence that these K-epoxides are not the major carcinogenic intermediates, at least for some hydrocarbons. Thus, Baird *et al*. have found that 7-methylbenz [a] anthracene and its K-epoxide form significantly different RNA-bound products on treatment with mouse embryo cells in culture (*Cancer Res.*, **36**, 2306; 1976), although

they were unable to actually characterise these products. This is of course not surprising, since the extent of *in vivo* binding is low—at most one hydrocarbon molecule per 10⁴ nucleotide residues. Recently devised techniques of extremely sensitive spectrofluorimetry have helped in this respect. Thus, Daudel *et al*. (*FEBS Lett.*, **57**, 250; 1975) have been able to compare directly spectra of DNA isolated from solution-binding with benzo [a] pyrene, with those from treated mouse skin as well as with spectra from several putative metabolite–DNA complexes. Their results clearly indicate that the K-region 4,5-epoxide is not the species

that reacts with DNA, and instead provide suggestive, though not conclusive evidence that a non-K-region diol-9,10-epoxide is primarily implicated in the metabolic activation of benzo[a]pyrene. Support for this has also come from the similar studies of Weinstein *et al.* (*Biochem. biophys. Res. Commun.*, **70**, 1172; 1976). Weinstein, Harvey and their collaborators have now succeeded in isolating and identifying the adducts from RNA degradation after *in vivo* incubation of benzo[a]pyrene in bovine bronchial explants. The intermediate is the (I) 7β,8α-dihydroxy-9α,10α-epoxy-7,8,9,10-tetrahydro derivative, that is a non-K-region diol epoxide, with subsequent attachment of the 10 position to the 2-amino guanine group (*Science*, **193**, 592; 1976). This compound has also been shown by Newbold and Brookes (*Nature*, **261**, 53; 1976) and Sachs *et al.* (*Proc. natn. Acad. Sci. U S.A.*, **73**, 607; 1976) to be by far the most active mutagen (and therefore carcinogen) of a range of possible benzo[a]pyrene metabolites, including the K-epoxide.

There still remain however, several unanswered questions concerning the ultimate carcinogenic intermediates arising from benzo[a]pyrene, and indeed from polycyclic aromatic hydrocarbons in general. Certainly the diol epoxide is by no means the only reactive intermediate; others remain to be isolated and defined. One should perhaps add that although there is now some evidence that the benzo[a]pyrene diol epoxide is an intermediate in other tissues and species during hydrocarbon binding to DNA, it is not really known whether the same is true for binding to other cellular macromolecules. □

Nucleic acid modification at Erlangen

from Ernest Borek

An international conference on post-transcriptional modification of nucleic acids, organised by Drs Helga and Walter Kersten, was held at the University of Erlangen on July 19–24, 1976.

THE field of nucleic acid modification is burgeoning, making this a most timely meeting. In eukaryotes every species of nucleic acid is modified and the number of known modifications is increasing—six new modifications were reported at this meeting.

W. Kersten (University of Erlangen)

led a round table discussion on the status of tRNA methyltransferases and isoaccepting tRNAs in tumour tissue. In all tumours, except benign ones, the tRNA methyltranferases are altered both qualitatively and quantitatively compared with their counterparts in normal tissue. All tumours examined contain novel isoaccepting tRNAs some of which are also found in embryonic tissue. There is no evidence for extensive hypermethylation of bulk tRNA isolated from tumour tissue, however, but since there are only a few tumour-specific isoaccepting tRNAs in each tumour, each of these purified tRNAs will have to be analysed. The extraordinarily high turnover of tRNAs in tumour tissue has been established recently by reliable analyses of urinary excretion products stemming from the breakdown of tRNA, determined at the US National Cancer Institute. A normal adult may be excreting about 2 mg of dimethylguanosine in his urine per day whereas some patients with Burkitt's lymphoma and cancer of the breast may be excreting as much as 20 mg. If we assume a tumour burden of 1–2 pounds per 100-pound patient, then 1–2% of his metabolising machinery may be producing the breakdown product at a rate 500% higher than a corresponding

Curious atom

from W. T. Toner

A HYDROGEN-like atom in which short-lived elementary particles substitute for both the nucleus and the orbital electron has been detected as the direct product of the decay of a third unstable elementary particle. Professor Schwartz and his colleagues (Coombes *et al.*, *Phys. Rev. Lett.*, **37**, 249; 1976) hope that a study of this most intriguing scientific curiosity will yield new information on the properties of π and μ mesons.

The parent is the longer lived of the two neutral K-mesons, the K_L' which frequently decays into a π meson, a μ meson and a neutrino. In a minute fraction ($\sim 10^{-7}$) of these decays, the oppositely charged mesons are emitted travelling by chance in the same direction with such similar velocities that they remain together long enough to be bound by their mutual electrical attraction. The bound systems are observed after they have passed down a long straight channel swept clear of charged particles by a magnetic field, in an apparatus which detects their charged components when the "π-μ atoms" are dissociated in an aluminium foil.

amount of normal tissue. Some unknown effect of the tumour on the metabolism of the rest of the body cannot be excluded. Such a high turnover of tRNA has also been demonstrated recently in an animal model.

The significance of these aberrations in tumour metabolism is, of course, obscure. They may be part of the oncogenic process or they may contribute to the maintenance of the morbid condition. In view of the many functions and activities of tRNA the problem presents an interesting challenge.

Nothing is yet known about methylation of DNA in tumour tissue. The technology of eukaryotic DNA methyltransferase is still lagging behind the rest of the field, because of the impossibility until recently of preparing soluble enzymes, and there are still conflicting reports of cofactor requirements.

Large strides have been made, however, in the understanding of rRNA modification and processing. J. Dahlberg (University of Wisconsin) identified some of the "spacer" polynucleotides between 16S and 23S rRNA genes as tRNA genes, two of which have been identified as tRNA^{Glu}₁ and tRNA^{Ile}₁. This distribution can account for the coordinate control of the transcription of tRNA and rRNA. R. J. Planta (University of Amsterdam) reported on the species specificity of rRNA modification. The ψ (pseudouridine) content of bacteria, yeast and human rRNA, per thousand nucleotides, is 2.5, 8 and 12 respectively. New modified nucleosides are still turning up. Planta has identified a ψ methylated at N_1 and hypermodified at N_3 by an α-amino-α-carboxyl-propyl moiety. (The latter stems from methionine by the elimination of the thiomethyl group) Nishimura had reported earlier on the modification of U in tRNA by the same fragment of methionine. The versatility of S-adenosylmethionine as the progenitor of modifications is astonishing.

Nishimura (National Cancer Institute, Tokyo) reported his latest studies on the fascinating hypermodified Q base which is present, adjacent to the anticodon, in several tRNAs of many species. What had been thought to be an isomer, Q*, is actually a still more modified form of it. In Q*, position 4 of the cyclopentene ring can be condensed with either mannose or galactose. The method of achieving modification of tRNA by the Q base is fascinatingly complex. In 1972, Farkas (University of Tennessee) discovered an enzyme which inserts guanine into tRNA^{His} of reticulocytes. Nishimura has now found a function for this "guanine insertase"—it introduces a guanine to replace the Q base

without cleavage of the phosphate linkage. Whether an enzyme to reverse the structure, "Q insertase", also exists is not yet known. The complexity of these reactions is significant in view of the known fluctuation in Q base content in *Drosophila* during metamorphosis, and the elevated content of Q* in tumour tissue.

H. Kersten (University of Erlangen) reported on the origin of the methyl groups in thymine of tRNA in Gram-positive organisms. In all but one out of some 20 tested the methyl group is synthesised at the macromolecular level by way of the folate pathway. In all other organisms, including mammals, that methyl group originates from S-adenosylmethionine. On the other hand, the methyl groups in thymine in the rRNA of Gram-positive organisms come from S-adenosylmethionine. It was pointed out that the folate pathway of synthesis must be the more primitive since this is the origin of the methyl group of thymine in DNA. Methylation by enzymes at the macromolecular level is probably a later development since it is the method used to imprint species individuality on DNA and in tRNA. An attempt to resolve the paradox of the two pathways for the methyl groups in tRNA and rRNA was attempted by W. Kersten by the suggestion that tRNA is more primitive than rRNA. This is an interesting suggestion for those interested in the evolution of biomacromolecules.

H. Aschoff and W. Kersten (Erlangen) reported on an 800-fold purification of tRNA 7-G transmethylase from *E. coli*. Starting with kilos of *E. coli* they obtained a protein of 300,000 daltons which produces only 7-meG in tRNA from *B. subtilis* in which a G in the extra arm is not methylated. This signal achievement 15 years after their discovery of the transmethylase highlights the difficulty inherent in purifying these complex enzymes.

One session was devoted to the modification of mRNA in eukaryotes. This area, too, is becoming increasingly complex. For example, D. T. Dubin (Rutgers University) has found a hitherto unknown "capping" of Sindbis virus-specific mRNA with an N₇-dimethyl-7-monomethyl G.

It is interesting to note the impact of policies of governance of science on the particular field covered by the meeting. Pre-eminence in the technology of tRNA modification has now passed from America to Germany and Japan, and one reason for this must lie in the pattern of the granting mechanism in the USA since no one in his right mind would dare tackle such problems on a 2-year NSF grant. There have been two international

conferences on nucleic acid modification in the past 2 years; the first one was convened by Italian colleagues and the current one by German workers. To neither of these conferences was a contributor from Great Britain invited, which may well be because the MRC decided a few years ago that the study of modification of nucleic acids should have a low priority for support. What a pity to exclude the people who initiated modern biochemistry from tackling one of the most challenging problems in the field. ☐

Mysterious meteorites

from M. G. Edmunds

Some forty meteorite chemists and astrophysicists met for an international four-day workshop on isotopic abundance anomalies organised by the Department of Applied Mathematics and Astronomy, University College, Cardiff, and held at Gregynog Hall on August 10–13, 1976.

THE early history of the Solar System has always provided ample scope for speculation, yet recent experimental investigations of the isotopic composition of certain meteorites have provided even more food for thought. A class of meteorites, the carbonaceous chondrites, has traditionally been regarded as the best available sample (except for some volatile elements) of the material out of which the Solar System formed. It came as rather a surprise, therefore, when work in the past few years showed that the relative isotopic abundances of particular elements in these meteorites are anomalous in that they can differ significantly from the so-called "cosmic" abundance ratios which are based on an amalgam of meteoritic, solar spectra and solar wind data. The workshop at Gregynog evolved into discussion of two major topics. The first problem was to sort out the experimental evidence to see exactly which elements really do show isotopic anomalies, and the second was to sift through possible theoretical mechanisms which could account for the origin of these anomalies. The three main possibilities examined were: chemical fractionation of isotopes occurring within some region of the early solar nebula out of which solid planets condensed; the production of particular isotopes in nuclear reactions caused by irradiation of the nebula by

Coated vesicles and clathrin

In a recent squib in these columns, Matus suggested that the protein molecule, "clathrin", which forms the lattice-like coats around coated vesicles (Pearse, *J. molec. Biol.*, 97, 93; 1975; *Proc. natn. Acad. Sci. U.S.A.*, 73, 1255; 1976; Crowther *et al.*, *J. molec. Biol.*, 103, 785; 1976) should be renamed "cytonexin". His reason for this is that E. G. Gray had described a fibrous network in presynaptic terminals which he called a cytonet (*J. Neurocytol.*, 1, 363; 1972; *Brain Res.*, 62, 329; 1973). At that time, Gray noted that the bristles on coated vesicles were indistinguishable (in the electron microscope) from the surrounding cytoplasmic matrix. As Matus states, Gray (*J. Neurocytol.*, 4, 315; 1975) has since repudiated the existence of this cytonet, concluding that it is an artefact of fixation probably arising from the precipitation of cytoplasmic proteins. Despite this, Matus now wishes to preserve this cobweb from the past and make us believe that the cytonet still exists, and that its constituent molecules can be identified by electron microscopy as clathrin, the molecule of which the striking polyhedral coats of coated vesicles are made.

B. M. F. PEARSE

M. S. BRETSCHER

A. I. MATUS REPLIES: Pearse and Bretscher complain about my rehabilitation of the cytonet on the grounds that Gray has repudiated it (*J. Neurocytol.*, 4, 315; 1975). However in that paper Gray ascribed both the cytonet and vesicular coats to a fixation artefact. If we are now to believe that vesicle coats are a genuine feature of the *in vivo* state, as Pearse's work suggests, then Gray's earlier description of the cytonet (*J. Neurocytol.*, 1, 363; 1972) is also unchallenged. Having seen many of Professor Gray's micrographs I have never doubted its existence. However, contrary to the suggestion made by Pearse and Bretscher I have no desire to make anyone believe in the cytonet. I do believe that it is important to point out an alternative to the restricted view afforded by their premature neologism for the coat/cytonet protein.

energetic particles; and the inclusion into meteorites of solid grains from outside the Solar System which had condensed in the stellar explosions which form elements. The grains could hold radioactive atoms and their decay products, carrying the isotopic anomalies from the nucleosynthesis sites.

It seems that definite anomalies have been demonstrated experimentally for several volatile elements. Anomalies for Ne and Xe are already well established. R. N. Clayton (University of Chicago) reported compelling evidence that enrichments of up to 5% in ^{16}O occur, with variations within a given meteorite implying a fine-scale, mineralogically-controlled heterogeneity. R. N. Clayton and J. Kerridge (University of California, Los Angeles) both presented arguments for ^{15}N anomalies, Kerridge concentrating on the fascinating possibility that the ^{15}N content of lunar soil decreases with decreasing age of the soil. Among the welter of possible mechanisms which were discussed for this correlation, the two most promising were time variations in either accretion of interstellar material rich in ^{15}N (for example from grains formed in nova explosions) or in the ^{15}N content of the solar wind incident on the moon and implanted into the soil.

An analysis of the meteorite Arapahoe was reported by F. Podosek (Washington University, St Louis), showing a $^{129}Xe/^{132}Xe$ ratio which is only about half that found in younger chondrites, the lunar soil or present-day solar wind. He argued that the solar nebula must have possessed some inhomogeneity, perhaps by locking up the radioactive precursor of ^{129}Xe in dust and forming Arapahoe out of a ^{129}Xe-poor gas region. D. D. Clayton (Rice University, Texas and University College, Cardiff) was quick to suggest that the ^{129}Xe-rich dust could have been formed in supernova explosions. This idea met considerable opposition, on the basis of the argument that supernova-condensed grains would be unlikely to give the perfectly normal (compared with other chondrites) pattern of release of iodine gas isotopes which is observed when Arapahoe is heated.

Although anomalies in volatiles seem established, the existence of anomalies in refractory elements is much more doubtful. D. N. Schramm (Enrico Fermi Institute, Chicago) reviewed the ^{26}Mg anomaly measurements, and he concluded that there was definite enrichment in addition to chemical fractionation effects. But this was the only refractory element for which an anomaly seemed to be unambiguously demonstrated. Essentially negative results were reported by F. Begemann (Max-Planck-Institut, Mainz) in searches for anomalies in Hg and K, and by L. Grossman (University of Chicago) for Os.

Everyone agreed that meteorites are rather inhomogeneous bodies. R. Lewis and B. Srinivasan (University of Chicago) gave accounts of the separation treatments they have used (with E. Anders) to find within meteorites those minerals which actually carry the anomalous isotopes. One mineral known as "Q", which has been isolated but not identified, contains up to 80% of the trapped noble gases and 50% of the anomalous Xe, although it comprises less than 0.1% of the meteoritic mass. Anders' Chicago group prefers the interpretation that a volatile superheavy element underwent fission to produce the observed Xe anomalies, rather than believing that the minerals carrying the anomalies represent solid material from outside the solar system. O. K Manuel (University of Missouri) repeatedly pointed out, however, that Xe anomalies are correlated with the trapped helium and neon content, implying that some kind of differentiation rather than fission has taken place. Evidence that fission of ^{244}Pu has produced some of the Xe is fairly strong. K. Marti (University of California, La Jolla), P Pellas (Orsay) and Podosek, presented observations of fission tracks, but the number of tracks indicates only a third of the fission required to give all the trapped Xe It is not clear whether the missing tracks have been destroyed, or if an alternative source of Xe is required.

Discussion also ranged over theoretical models for the condensation of solids, both during the formation of the planetary system and in explosive nucleosynthesis events. Grossman suggested that a model for the solar nebula in which the temperature never rose high enough to evaporate all pre-existing solids was at least not inconceivable, and T. Gold (Cornell) pointed out that such conditions could well hold in a slowly contracting disk of material about the Sun. Gold also emphasised that a disk could act as a "flypaper" to collect interstellar material, and that if the disk lasted for an appreciable time then "extra-solar" rather than "pre-solar" solid material could well be captured. M. Newman (Caltech) reported models for accretion of interstellar matter by the Sun, but such effects were likely to be important only if the solar nebula had passed through a dense interstellar cloud. P. M. Solomon (Stony Brook, New York and University College, Cardiff) reviewed the evidence for a large part of interstellar matter existing in the form of a giant dense molecular cloud. and N. C. Wickramasinghe (University College, Cardiff) reviewed data and theories relating to the dust component of interstellar clouds. If a large dense cloud were the nursery of the Solar System, then accretion effects could be important, and be highly relevant to composition of meteoritic matter.

Condensation sequences within the different nuclear zones expanded from a supernova explosion were described by both J. Lattimer (Enrico Fermi Institute, Chicago) and S. Ramadurai (University College, Cardiff). Such equilibrium calculations were recognised as being necessarily rather unrealistic, but did suggest some minerals which might be looked for in meteorites as signatures of dust formed in explosive events. It was also recognised that the mixing between supernova synthesis zones remains an important, but unknown, process in determining the chemical composition of solid products.

Another possible origin for the isotopic anomalies was briefly discussed, the suggestion that a local supernova event at the time of formation of the Solar System sprinkled the isotopes or their radioactive precursors in some inhomogeneous way over the condensing solar nebula. Even the proponents agreed that a model involving nuclear reactions from particle irradiation in the solar nebula ran into extreme energetic difficulties, both in the total amount of energy required and the convoluted shape of the energy spectrum necessary to avoid overproduction of certain species. Although an extensive reaction network reported by S. Woosley (University of California, Santa Cruz) could reproduce some of the anomalies, it certainly could not (for example) explain that of ^{16}O.

More experimental data should soon be available. The elegant experimental techniques of very small scale analysis described by M. Maurette (Orsay) and his colleagues, showing substructure on the scale of 100 Å in 1-μm-sized inclusions, will be able to give chemical and mineralogical information on individual inclusions in the meteorites. Whether real anomalies exist for any non-volatile elements other than Mg remains to be seen, but the results of searches for Si, Ca, Ti, S and Fe are eagerly awaited.

The workshop provided a very enjoyable informal environment for the exchange of ideas. The best interpretation of the anomalies remains unclear—D. C. Black's (*Geochim. cosmochim. Acta,* **36**, 377; 1972) original suggestion of grains from nucleosynthesis sites has its attractions, and has been enthusiastically championed by D. D Clayton and F. Hoyle. But this model still has some serious difficulties, and a lot of detailed mechanism to be worked out At any rate, it looks as if the isotopic anomalies still have much to tell us about how the Solar System formed. □

review article

Electrochemical, solid state, photochemical and technological aspects of photoelectrochemical energy converters

Joost Manassen, David Cahen & Gary Hodes*

Avraham Sofer†

Different aspects of photoelectrochemical energy conversion are discussed. It is shown, by conservative estimates of the energy losses in such a system, that the available photopotential is smaller than the optical bandgap of the semiconductor by at least 1 eV. It is shown that water decomposition using only one photosensitive electrode is energetically unfavourable and that, even when two photosensitive electrodes, are used, the production of electrical current is to be preferred over that of hydrogen.

THE idea of converting solar radiation directly into electrical energy by using electrochemical cells, is gaining in popularity[1-9]. It is a multidisciplinary field, which covers electrochemistry, solid state physics and photochemistry as well as technology and economics[4-9].

From reviewing the publications in these fields it is clear that specialists in one field often are not aware of the implications and restrictions of another field. It is therefore our intention to discuss in this paper some simple general principles rather than to discuss the different aspects in depth as is done in the specialised texts.

Electrochemical aspects

If an electrode is immersed into an electrolyte solution containing a redox couple R/O and is held at a certain potential, two reactions occur at the electrode surface:

(1) The cathodic reaction $O + ne \rightarrow R$ (e = electron, n = reaction valence)

and

(2) The anodic reaction $R \rightarrow ne + O$

If both reactions are rapid and the electrode is held at the equilibrium potential of the couple, a situation is established where the concentrations of R and O are equal. At more anodic potentials, O will be produced, and vice versa. In the case of a slow anodic reaction, for kinetic reasons, however, an anodic overpotential has to be applied to produce O, and the same applies to R when the cathodic reaction is slow. In cases where the cathodic and anodic reactions are one-step processes, we may expect the same kinetic barrier to exist for both reactions and both to be retarded accordingly, but this need not necessarily be so for a multi-stage reaction (see section on energy conversion).

In a photoelectrochemical energy converter, a photoactive electrode is connected with a counter electrode, which should

be as rapid as possible towards the reaction reverse to the one occurring at the photoactive electrode. If the photoactive electrode is illuminated, an electrical current may flow for two quite distinct reasons: one is photocatalysis and the other energy conversion.

Photocatalysis

Photocatalysis can occur only when the system is in a state such that current flow is thermodynamically possible, but is prevented by the kinetics. Such a thermodynamic state may be created, for example, if both electrodes are held at a different potential either by means of an external power source or by keeping anolyte and catholyte at a different composition or temperature. Current can be negligibly small under these thermodynamically favourable conditions, for example because of an overpotential at the photoactive electrode for the redox reaction. The kinetic barrier may be lowered by illumination leading to an increase in current flow. This current flows however because of photocatalysis of an electrochemical reaction and without conversion of light into electrical energy. Another possibility is that the presence in the solution of a material which can react chemically with R or with O, but fails to do so for kinetic reasons. Again, illumination may lower the kinetic barrier and because R or O is consumed, an electrical current will start to flow. In this case we have a photocatalysed fuel cell and again no conversion of light into electrical energy.

Energy conversion

Energy conversion can occur only when illumination changes the thermodynamics of the system, and not just the kinetics as in the case of photocatalysis. In photocatalysis we lower a kinetic barrier and get a decrease of overpotential. In energy conversion we change the energy level of one of the components of the reacting system and must get an underpotential. This means that illumination leads to a situation which in the dark is thermodynamically unfavourable. There is an additional

*Address: The Weizmann Institute of Science, Rehovot, Israel.

†Address: Chemistry Department Nuclear Research Center, Negev, Beer Sheva, Israel.

Table 1 Currents, potentials and power-output of a TiO_2 photo-electrochemical cell, as a function of the counter electrode

Counter electrode	Short-circuit current (μA)	Open-circuit potential (mV)	Maximum power (μW)
O_2 bubbling through solution			
Platinum on high surface carbon black	640	760	200
Platinised platinum	640	760	200
Smooth platinum	530	670	143
Without bubbling			
Platinum on high surface carbon black	640	760	200
Platinised platinum	580	720	175
Smooth platinum	460	640	120

Electrolyte: 1 N KOH. Light source: 150 W tungsten/Iodide. (Intensity equivalent to sunlight incident on the cell.) TiO_2 polycrystalline. Surface: TiO_2 electrode, 0.75 cm²; counter electrode, 4 cm².

condition, however. Imagine that illumination causes an anodic underpotential, which means that O is produced at potentials lower than the equilibrium potential. If the cathodic reaction is rapid, O will be reduced immediately back to R and no current will flow to the counter electrode. The additional condition is, therefore, a slow back reaction at the photoactive electrode.

An attractive system is the redox couple O_2/H_2O, which has been much studied in connection with the photoactive TiO_2 electrode[1–3,6,7–11]. If a TiO_2 electrode is illuminated with light, having an energy larger than its optical bandgap (see below), oxygen is evolved at potentials of more than a volt lower than the thermodynamic oxygen evolution potential. This does not mean however, that in a photoelectrochemical cell with a TiO_2 anode, hydrogen is evolved at the counter electrode. According to the principles described here, oxygen, if present, will be reduced back to water at the counter electrode. Electrodes, which have a low overpotential for oxygen reduction are known from fuel cell technology and in Table 1 (J.M., G H. and D.C., unpublished results) we show how the performance of such a photoelectrochemical cell (PEC) can be improved by changing the counter electrode.

The slow back reaction is ensured in this system because TiO_2 shows a large overpotential for oxygen reduction.

Hydrogen evolution has only been proven in this system by either giving a cathodic bias[7,10] or by having catholyte and anolyte of different composition[7,11]. In both cases a clear distinction has to be made between energy conversion with the assistance of light, and energy conversion solely by light.

In the next section we shall show that to maximise efficiency, overpotentials should be minimised. At the counter electrode this can be achieved by the use of electrocatalysts and a high surface area, as we have seen. At the photoactive electrode we demand a slow back reaction, however, and we have seen that for many cases this means that the forward reaction will also be slow (leading to a sizeable overpotential). An ideal case would be a multistage redox couple, like the O_2/H_2O system, where in principle the forward reaction (oxygen evolution) may be fast and the backreaction (oxygen reduction) may be slow. This, however, is an exceptional case and redox reactions at semiconductor electrodes tend in any case to be slower than at metals. The conditions of a slow back reaction at the photoactive electrode will therefore, in most cases mean that the overpotential of the forward reaction will be appreciable.

Solid state aspects

For an understanding of the origin of the photopotential generated in a photoelectrochemical cell and to estimate its magnitude, it is necessary to understand what is happening in a semiconductor when it is brought into contact with an electrolyte solution and subsequently illuminated. The processes involved can be illustrated schematically by plots of electron energy levels across the semiconductor–electrolyte interface.

Fig. 1a shows the energy levels of an n-type semiconductor in contact with vacuum. When the semiconductor is immersed in an electrolyte solution, containing a redox couple with redox potential E_{redox}, the situation as depicted in Fig. 1b will result, after thermodynamic equilibrium between the two phases is established (that is, the electrochemical potential in both phases is equal). [To put the redox potential on the same energy scale as that used for the semiconductor, both can be referred to the energy level of an electron *in vacuo* at infinity (E_{vac}). It has been shown that E_e (redox) in volts $= E_e$ (redox) in eV + 4.5 eV, that is the energy level for the H^+/H_2 redox couple is 4.5 eV below[1a] that of E_{vac} (4.3 eV according to ref. 13). We assume here, for simplicity, that surface states do not have a role in determining the energetic picture of the electrolyte-semiconductor interface, that is, we deal with the so-called Schottky case. If they do play a part, they can determine the band bending to a greater or lesser extent, independent of the electron affinity of the semiconductor.] This equilibrium will be reached by the passing of majority carriers (electrons in this case) from the semiconductor to the solution, causing a redistribution of positive and negative charge carriers at and near the surface until the Fermi level of the solid equals the redox potential of the solution. Schematically this is shown by a bending of the
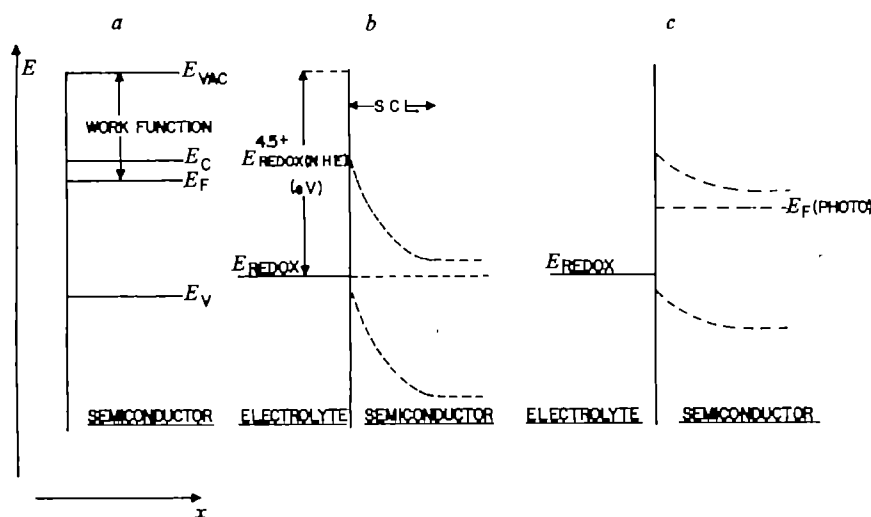
Fig. 1 Plots of energy levels: *a*, n-type semiconductor *in vacuo*; *b*, n-type semiconductor/electrolyte in the dark; *c*, n-type semiconductor/electrolyte, when illuminated. SCL: space charge layer; E_c: bottom of conduction band; E_v: top of valence band; E_{vac}: energy level of electron in vacuum; E_F: Fermi level; $E_{(redox)}$: Redox potential of electrolyte; NHE: Normal hydrogen electrode; E_F(photo) Fermi level of illuminated electrode.

valence and conduction bands towards the semiconductor surface, forming the space charge layer.

Illumination by radiation with energy content more than E_c-E_v $(= E_g$, the optical bandgap) leads to charge creation and because of the potential difference across the space charge layer, majority carriers (electrons in Fig. 1c) move towards the bulk of the semiconductor electrode and minority carriers (holes in Fig. 1c) towards the electrode surface. This movement counteracts the band bending (that is, causes a partial return to the charge carrier distribution of Fig. 1a) and creates a new situation in which the semiconductor Fermi level is no longer equal to the redox potential of the electrolyte solution. If now the semiconductor electrode is connected with a counter-electrode having the properties defined in the previous section, a photopotential V is established between the two electrodes, which can reach a maximum value of $E_F(\text{photo})-E_{redox}$ (Fig. 2).
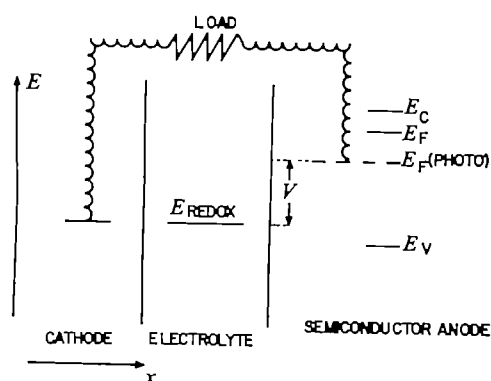
Fig. 2 Schematic illustration of the photopotential (V) obtained in a photoelectrochemical cell. The relative positions of E_v, E_F, $E_F(\text{photo})$ and E_v, from Fig. 1 are indicated.

In a working, corrosion-free cell, several kinds of energy losses occur which can be estimated, using Fig. 2 as a guide. The energy balance can be composed of the following units

$$hv = E_c-E_v = (E_{redox}-E_v)+(E_F-E_F[\text{photo}])+(E_c-E_F)+ iR+\eta_c+\eta_p+V \tag{1}$$

where iR are resistance losses in the system, η_c and η_p are overpotentials at the counter electrode and the photoelectrode, respectively.

We shall now make conservative numerical estimates of the different factors involved:

The difference in energy between the potential of the redox couple and the valence band, $(E_{redox}-E_v)$, has to be chosen equal to or larger than λ. Here λ is the reorganisation energy, which has been discussed extensively elsewhere[14-16]. It has to be introduced, because when $(E_{redox}-E_v)<\lambda$ the possibility exists of minority-carrier (hole, in this case)-injection from the redox couple into the semiconductor valence band, which runs counter to the effect we are looking for and may also cause corrosion of the semiconductor.

The value of λ has been estimated variously between 0.4 and 0.8 eV[16,17] and we shall take it here at 0.4 eV.

The difference in energy between the Fermi level *in vacuo* and the Fermi level in contact with electrolyte under illumination, $(E_F-E_F[\text{photo}])$, has to have a non-zero value in order for charge separation to occur. Its allowed value will depend on the width of the space charge layer and the lifetime and mobilities of the charge carriers. To get practical currents, its value may be conservatively estimated at 0.1 eV.

The difference in energy between the conduction band and the Fermi level, (E_c-E_F), may be small, but may also be several tenths of an eV (for TiO_2 it may vary between 0.003 and 0.1 eV (ref. 18)).

The terms iR and η_c can be minimised by practical cell construction and we estimate their combined values, together with that of (E_c-E_F) to be 0.1 eV, which is a conservative estimate indeed.

We have shown in the preceding section that η_p cannot be too small, because of the demand of a slow back reaction and because redox reactions at semiconductor electrodes tend to be slow. Therefore η_p has to be given a larger value and we put it at 0.4 eV.

By substituting these estimates in formula (1), we get

$$hv = E_c-E_v = V+1 \text{ eV} \tag{2}$$

In other words, the difference between the obtainable photopotential and the optical bandgap of the semiconductor is at least 1 eV in a working photoelectrochemical cell.

This conclusion has some important consequences. (1) If we want to use a semiconductor with an optical bandgap, so as to give optimal utilisation of solar radiation (~ 1.4 eV: ref. 17), the maximum photopotential attainable will be ~ 0.4 eV. The maximum energy obtainable from solar radiation is about 100 mW cm^{-2}. A PEC converting 10% of this energy would produce 10 mW cm^{-2}, which at 0.4 V means a current of ~ 25 mA cm^{-2}. Such a current makes our estimates for the iR and overpotential losses somewhat optimistic. (2) If our aim is to decompose water, V has to be >1.23 eV and the optical bandgap of the semiconductor electrode >2.23 eV. Therefore to decompose water into H_2 and O_2 we have to accept the loss of an important part of the solar spectrum ($\lambda > 550$ nm). It is possible to get around this problem by using both an n-type semiconductor photoanode and a p-type semiconductor photocathode, in which case two photopotentials, V_c and V_a (see Fig. 3) are obtained, which are additive[10].
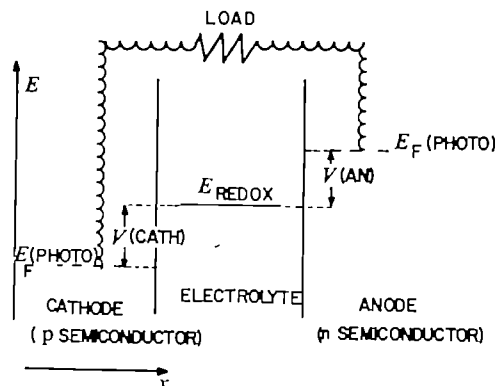
Fig. 3 A schematic illustration of the photopotential obtainable when a combination of a photoactive n-type anode and a photoactive p-type cathode is used.

It is interesting that the water decomposition reaction, which occurs naturally in green plant and algal photosynthesis using chlorophyll, also requires two light quanta to obtain the necessary electrochemical potential difference, and that of the light input of 1.8 eV of each quantum, only 1 eV is utilised for the generation of an electrochemical potential difference[11].

Most of the photoelectrochemical cell systems described are complicated by corrosion problems and its prevention is one of the most difficult obstacles to practical realisation of this system. The Cd-chalcogenide based cell, in which the corrosion problem is circumvented by the use of the S/S^{2-} redox couple, shows great promise in this respect[12,13].

Corrosion

It is clear that the electrodes and the redox couple must be entirely stable in the environment of the electrolyte used. There

is however an additional phenomenon, that of corrosion under the influence of light. With CdS the S^{2-} ions are oxidised to sulphur and with ZnO the O^{2-} ions are oxidised to oxygen and the Zn^{2+} ions pass into the solution[24,25]. This means that both photoactive anodes are not stable towards the strong oxidising conditions which exist at their surfaces during illumination. On the other hand, TiO_2 has been shown to be stable under these conditions, as has WO_3 (ref. 26) and Ta_2O_5 (ref. 27). Both cadmium and zinc have only one stable valence state, whereas titanium, tungsten and tantalum are stable in several valence states. It may be, therefore, that one way to obtain stability against photocorrosion in these conditions is by the presence of ions in the semiconductor which can accommodate several valence states.

Photochemical aspects

Problems connected with the heterogeneity of solar radiation have been extensively discussed elsewhere and need not be dealt with here[19].

One has to realise however, that for reasonable conversion efficiency, the light has to be absorbed within the space charge layer, where the necessary potential gradient exists for charge separation to occur. For a material having a decadic absorption coefficient α, 90% of the incident radiation will be absorbed in a layer of thickness

$$L = 1/\alpha \quad (\alpha \text{ in cm}^{-1})$$

A value for α of $10^4 - 10^5 \text{ cm}^{-1}$ is typical of many semiconductors in the strongly absorbing range (both TiO_2 and Ta_2O_5 have exceptionally high absorption coefficients close to 10^6 cm^{-1}) and this would demand a thickness of the space charge layer of about $10^2 - 10^4$ Å. To minimise ohmic losses, we should like to make the semiconductor electrode as conductive as possible. As the space charge layer narrows with increasing conductivity, it can easily reach a stage, at which most of the light passes the space charge layer, with a resultant decrease in conversion efficiency.

Technological and economic aspects

The terms photoelectric energy storage and photoelectric energy conversion are sometimes confused.

A photoelectric energy converter produces electrical energy under illumination and does not store it unless specific storage electrodes can be used[22]. As long as energy production by such a device is small, compared with the overall use of electricity, the electrical current produced can be fed directly into the existing distribution system. This will save fuel, but the capacity of alternative generating capacity will have to remain the same to prevent serious shortages on cloudy days. With the increasing emphasis on the use of nuclear energy, fuel saving will become less important, since fuel costs in a nuclear power station are much smaller than those in a conventional station. Better and cheaper means of storing electricity, however, remain desirable, and hence the practical importance of photoelectrochemical decomposition of water into hydrogen and oxygen.

We have shown that to achieve this with the use of only the photoactive electrode is very inefficient, (no examples have as yet been given in the literature to show that this is feasible without adding energy to the system[7,10,11]). It must be kept in mind that if hydrogen can be produced by the use of a photoactive anode and a reversible counter electrode, most of the energy from hydrogen production can be converted into electrical energy by the use of an oxygen cathode. Fuel cell technology has produced oxygen cathodes with low overpotentials for oxygen reduction under the conditions of low current density which invariably will exist in a solar energy device which does not concentrate the sunlight.

Electrical current may be transported to a central electrolysis plant, which converts water into hydrogen and oxygen at a high efficiency (90% has been claimed[28]). This has to be carefully weighed against the intrinsic lower efficiency of hydrogen production and the appreciable technological difficulties in transporting small quantities of hydrogen (70 ml of hydrogen per (mA cm^{-2}) per min per m^2 of electrode area) leakfree to a central container. If a combination of an n-type and p-type semiconductor is used, both have to exposed to sunlight. This is only attractive when both photoactive electrodes work at about the same efficiency. If one of the two is appreciably more efficient, it is again more favourable to use the most efficient electrode for electricity production alone and to produce the hydrogen in a central electrolysis plant. The only way out of this dilemma would be to use one electrode which absorbs one part of the solar spectrum and positioning it in front of another electrode, which responds to the radiation which has passed the first one.

It seems therefore that the technological prospects of photoelectrochemical decomposition of water are rather dim because there is always the alternative of producing electrical current instead at a higher efficiency and using the current for electrolysis of water in a central plant.

Advantages of photoelectrochemical cells

In the above discussion several inherent disadvantages of the photoelectrochemical cell have been emphasised, which may be less severe for dry photovoltaic devices. There are several arguments however, in favour of pursuing research and development of the photoelectrochemical cell. It has been shown that high efficiency can be obtained, using a thin polycrystalline semiconductor surface. This is because of the much better contact obtained between a polycrystalline solid and a liquid than is possible between two polycrystalline solids, and the "single crystal barrier", which is often used as an argument against cheap solid-state solar cells, does not have to be overcome here. Therefore the production cost of a photoelectrochemical cell can be considerably less than that of dry cells. "Cost" does not only mean the money to be paid, but also the energy-input in production. By mass-production techniques it will certainly be possible to decrease the cost in terms of money of dry photovoltaic cells, but it is much more difficult to bring down the cost in terms of energy and it is in this respect that the photoelectrochemical cell has a chance to compete with dry photovoltaic devices.

1 Fujishima, A., and Honda, K., *Nature*, 238, 37–38 (1972).
2 Gerischer, H., *J. electroanal. interfacial Electrochem.* 58, 263–274 (1975)
3 Archer, M. D., *J. appl. Electrochem.*, 5, 17–38 (1975).
4 Chu, Y. G., and Anderson, W. W., *Appl. phys. Lett.*, 27, 183–184 (1975)
5 Kooney, J., Wormstein, D. H., and Hass, G. M., *Nature*, 253, 719–720 (1975).
6 Jayadevaiah, T. S., *Appl. phys. Lett.*, 25, 399–400 (1974)
7 Nozik, A. J., *Nature*, 257, 383–386 (1975).
8 Hardee, K. L., and Bard, A. J., *J. electrochem. Soc.*, 122, 739–742 (1975).
9 Mollers, F., Tolle, H. J., and Memming, R., *J. electrochem Soc.*, 121, 1160–1167 (1974).
10 Fujishima, A., Kobayakawa, K., and Honda, K., *Bull. chem. Soc. Japan*, 48, 1041–1042 (1975).
11 Wrighton, M. S., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 72, 1518–1522 (1975).
12 Lohmann, F., *Z. Naturforschung*, 22A, 843–44 (1967).
13 Trasatti, S., *J. electroanal. interfacial Electrochem.*, 52, 313–329 (1974).
14 Levich, V. G., in *Advances in Electrochemistry and Electrochemical Engineering*, 4 (edit. by Delahay, P.), 249–371 (Interscience, New York, 1966).
15 Gerischer, H., *Surface Science*, 18, 97–122 (1969).
16 Gomes, W. P., and Cardon, F., *Z. Physik. Chem. Neue Folge*, 86, 330–334 (1973).
17 Gerischer, H., *Faraday Disc., Chem. Soc.* No 58, 219–236 (1974).
18 Goodenough, J. B., in *Progress in Solid State Chemistry*, 5, (edit. by Reiss, H.) 145–399 (Pergamon, New York, 1971).
19 Fischer, H., *Festkorperprobleme*, 14, 153–182 (1974).
20 Yoneyama, H., Sakamoto, H., and Tamura, H., *Electrochimica Acta*, 20, 341–345 (1975).
21 Kok, B., in *Proceedings of the NSF-RANN Workshop on Bio-Solar Conversion* (edit. by San Pietro, A.) 22–30 (National Science Foundation, Bethesda, Maryland, 1973).
22 Hodes, G., Manassen, J., and Cahen, D., *Bull. Isr. Phys. Soc.*, 22, 100 (1976); *Nature*, 261, 403–404 (1976).
23 Ellis, A. B., Kaiser, S W., and Wrighton, M. S., *J. Am. chem. Soc.*, 98, 1635–1637 (1976).
24 Gerischer, H., and Mayer, M., *Z. Physik. Chem. Neue Folge*, 74, 302–318 (1971).
25 Morrison, S R., and Freund, T., *J. chem. Phys.*, 47, 1543–1551 (1967).
26 Hodes, G., Cahen, D., and Manassen, J., *Nature*, 260, 312–313 (1976).
27 Van Gool, W. Ch., Pistorius, C. A., and Winkel, P., *Phillips Res. Repts.*, 13, 265–276 (1958).
28 Ramakumar, R., Allison, H. J., and Hughes, W. L., *Solar Energy*, 16, 107–115 (1974).

articles

γ-ray bursts from thermonuclear explosions on neutron stars

S. E. Woosley & Ronald E. Taam

Lick Observatory, Board of Studies in Astronomy and Astrophysics, University of California, Santa Cruz, California 95064

It is proposed that γ-ray bursts originate from carbon detonations initiated by the accretion of matter on to the surface of a neutron star. The observations are interpreted in terms of this theory. Possible implications for the nuclear powered model of giant X-ray pulses are discussed briefly.

More than three years have now elapsed since the first published report of cosmic γ-ray bursts by Klebesadel *et al.*[1], but in spite of continuous monitoring and observation[2,3] the astrophysical origin of these phenomena still remains enigmatic[4]. Briefly, the dominant characteristics of the observed bursts are:

(1) Total durations of ~ 1–10 s with significant time structure on scales as short as 16 ms.

(2) Integrated burst energies at the earth of 10^{-5}–10^{-4} erg cm^{-2} for events which occur ~ 10 times per yr, with some evidence for more frequent less energetic bursts of ~ 10^{-7} erg cm^{-2} occurring perhaps several times per d (refs 3, 5).

(3) An energy spectrum that can be represented by an exponential from 100 to 400 keV tangent to a power law at higher energies[3]. The spectrum seems typical of a cooling black body of a temperature ~ 2×10^9 K (ref 6).

(4) The distribution of energetic sources seems to be crudely isotropic and uncorrelated with the observed strong X-ray sources, although Cyg X-1 is a possible exception[3].

The short time scale structure of these bursts suggests that they originate from a compact object[7], and the spectrum indicates a possible thermal source. We propose that the observed bursts are produced on the surfaces of neutron stars as a result of violent thermonuclear runaways initiated in accreted nuclear fuels. Attention is focused on the onset of carbon burning which we find, for reasonable choices of parameters, occurs in conditions that are very degenerate and potentially quite explosive. These same conditions have previously been found to result in the detonation of carbon in the core of highly evolved stars and have been investigated as a possible mechanism for the supernova phenomenon[8-13]. Such detonations are characterised by very short time scales and multibillion degree temperatures. Here, by using the gravitational potential well of a neutron star, we find that similar behaviour can occur for critical carbon masses of only 10^{22} g as opposed to the $1.4M_\odot$ usually required for a detonation initiated by self-gravitation.

The model

Consider a $1.4M_\odot$ neutron star of radius 10 km accreting hydrogen and helium at $10^{-10}M_\odot$ yr^{-1}. An accretion rate of this magnitude might be realised by a Roche lobe overflow of a less massive main sequence companion or perhaps by a single neutron star in a very dense cloud or nebula. This object would not appear as a strong X-ray source, and, because of

the softness of the resulting spectrum[14], might be difficult to detect at a distance of, say, 1 kpc. It will turn out that accretion rates of roughly this magnitude are required since the proposed detonation does not seem likely for markedly different values.

If the host neutron star has a strong magnetic field, the material is not accreted in a spherically symmetric manner, but instead accumulates in polar caps of ~ 10^{10} cm^2 or ~ 0.1 % of the surface area[4,15]. Initially any nuclear reactions that occur in this matter act only to strip down nuclei and fragment them. As additional material is accreted, however, compression occurs and nuclear fusion reactions, enhanced by electron screening, commence. Hydrogen burns to helium[15,16] and helium to carbon and oxygen[16]. These burning phases may become unstable, but, whatever the outcome, it is clear that little material escapes the gravitational field of the neutron star. Hansen and Van Horn[16] have found that the density of the helium-burning shell on a neutron star is ~ 10^6 g cm^{-3}. Thus one expects that the carbon formed by this burning will be relativistically degenerate. Because of the high conductivity of degenerate matter, the carbon will be maintained at a temperature close to that of the helium-burning shell. For the assumed accretion rate and geometry one can use the Hansen and Van Horn calculations of nuclear fusion on neutron stars to obtain an estimate for this temperature of ~ 4×10^8 K (Note that, in terms of accretion per unit area, $\dot{M} = 10^{-10}M_\odot$ yr^{-1} on the poles is equivalent to a uniform global rate of $\dot{M} \approx 5 \times 10^{-8}M_\odot$ yr^{-1}.)

It is of some interest to examine the state of this carbon layer after 4×10^{22} g of matter has accumulated: the density at the base of the carbon shell will then be ρ_b ~ 2×10^9 g cm^{-3} (calculated from the equation of hydrostatic equilibrium with pressure support from degenerate electrons). This shell of carbon, ~ 50 m thick, is capped by 1 or 2-m layers of hydrogen and helium containing ~ 10^{14} and ~ 10^{17} g respectively, both with active burning shells. Before reaching this density, the carbon layer is cooled by neutrino emission, with maximal losses in the regions of highest density. A positive temperature gradient will develop, so that conduction acts to heat the dense carbon rather than cool it. Once a density of ~ 10^9 g cm^{-3} is reached, however, the plasma neutrino losses become rather suddenly less effective in cooling, and at a density ~ 2×10^9 g cm^{-3} the nuclear energy generation from carbon burning dominates the neutrino losses. Any increase in density above this value leads to a nuclear runaway[8,9,17]. Because of the high degree of degeneracy, low specific heat, and extreme temperature sensitivity of the carbon-burning reaction the runaway will be quite violent. In the carbon-detonation supernova model, carbon ignites at temperatures and densities almost identical to those described here, and one achieves temperatures > 5×10^9 K and nuclear burning time scales (once a significant amount of carbon has burned) of ~ 1 ms (ref. 8). Not only does carbon burn, but also oxygen and their products as well. Here one

expects the entire layer of carbon to burn on a very short time scale either through a series of carbon flashes[13] or by the formation and propagation of a detonation wave[16]. If a detonation wave propagates the relevant time scale is < 1 ms. Ultimately, species in the iron group are produced and 7.6×10^{17} erg g^{-1} of nuclear energy is liberated. Thus in the case of the model neutron star discussed here, complete combustion of all available carbon releases 3×10^{40} erg. This energy is dissipated through neutrino processes, conduction, convection, and the expansion of the neutron star atmosphere. But unlike the case of a supernova, which it otherwise so closely resembles, the expansion that results from an explosion on a neutron star is negligible. The total energy generated by the explosion is only $\sim 0.5\%$ of the gravitational binding energy of the carbon. Moreover, any energy that goes into expansion is immediately regained on a time scale $\sim 10^{-3}$ s as the material is pulled back by gravity. Thus, the only immediate result of a carbon detonation in this case is to heat material to a very high temperature.

The relevant time scales at this point become very difficult to estimate especially without a detailed numerical model. The incandescent carbon will cool by emitting neutrinos, by conduction to the core and surface, and by convection. The conductivity of the hot carbon ($T \sim 5 \times 10^9$ K) is not as high as it was before the explosion (since $K_{cond} \propto T^2$). We estimate the thermal relaxation time scale for the 50 m of carbon to be $\sim 10^3 - 10^4$ s if cooling occurs by conduction alone, but this is an overestimate since convection will undoubtedly have an important role. Even with a detailed numerical model the physics of convection in a relativistically degenerate electron gas permeated by a 10^{12}-gauss magnetic field would be very uncertain. All one can say at this point is that a cooling time scale of a few seconds does not seem unreasonable.

The emission from the stellar surface may also have an unusual character. As mentioned previously, the layers of hydrogen and helium capping the carbon are very thin, and contain only $\sim 10^{-4}$ of the mass of the exploding carbon. It appears very likely that instabilities and convective overshoot will develop so that hot blobs of burning carbon will break through the surface of the neutron star. There incandescent blobs having temperatures of several billion degrees and dimensions perhaps of several metres will cool by emitting neutrinos and γ rays in times < 1 s. Additional temporal structure may result if the neutron star is an oblique rotator with rotation period less than the cooling time.

Comparison with observations

How do these results compare with the observations? The spectrum, following the above discussion, is that of a rapidly cooling multibillion degree black body, and the short time scale structure in the observed bursts relates to the cooling times of these small hot convective blobs. If one assumes that only 10% of the available energy is radiated in this form, the total photon energy is $\sim 5 \times 10^{38}$ erg. At a distance of 1 kpc this would give rise to an integrated flux at the Earth of $\sim 3 \times 10^{-5}$ erg cm^{-2} in good agreement with observations.

How many such objects would exist within 1 kpc of the Earth? Various arguments involving nucleosynthesis, supernova rates, and pulsar statistics suggest that there are $\sim 3 \times 10^8$ neutron stars in the entire Galaxy. If these are distributed uniformly throughout the Galaxy with a scale height $\lesssim 1$ kpc above the galactic plane one can expect $\sim 10^6$ neutron stars within 1 kpc of the Earth. Of these, perhaps 1% are in binaries. What fraction of the binaries are currently accreting mass at the required rate? This is connected with the mode of mass transfer which could be either a stellar wind, or Roche lobe overflow. If the former is the case, the mass transfer rate from the companion must be $\gtrsim 100$ times the accretion rate, since the neutron star subtends a small solid angle. This suggests that the mass loss rate from the companion is $\sim 10^{-8} M_\odot$ yr^{-1} which is typical for supergiant stars ($M \gtrsim 20 M_\odot$). However the duration of this mass transfer phase is < 10^5 yr (ref. 18). One then expects $10^6 \times 0.01 \times 10^{-5}$ < 1 object at this phase

of evolution. Clearly this is not in accord with the observations. If one considers Roche lobe overflow, a mass loss rate of $10^{-10} M_\odot$ yr^{-1} might be typical for a main sequence star less massive than the neutron star transferring mass on a nuclear time scale. One would then expect $10^6 \times 0.01 \times (0.01-0.1) \sim (100-1,000)$ objects. If the neutron star was formed by a type II supernova, there would be some doubt as to whether the 'desired' system would remain bound. For progenitor masses $\gtrsim 4 M_\odot$ (believed to be a probable lower limit) the binary would probably disrupt since the amount of mass loss in the supernova explosion would be > half the total mass of the initial system[19]. These arguments preclude the existence of γ-ray burst sources as evolving from type II supernovae. It is, however, not implausible that the systems could have evolved from type I supernovae (as, for example, may occur in cataclysmic variable systems)[20]. Accepting the above conclusion, one might expect the sources to be confined to the galactic plane, since the progenitors of the type I supernovae are believed to be old stars in the stellar disk. After the supernova explosion, however, there would probably be a random distribution of systems reflecting the randomness in the initial orientation of the binary orbits. Since these neutron stars could have been given high peculiar velocities and were produced as long as a billion years ago, a scale height of 1 kpc is not unreasonable. Thus we conclude, admittedly with a great deal of uncertainty, that 100 such objects could currently be active within 1 kpc of the Earth. Since a critical mass of carbon will accrete every year (giving rise to the possibility of repeated bursts from the same location) the number of events with an integrated flux in the range $10^{-5}-10^{-4}$ erg cm^{-2} would be ~ 100 yr^{-1}. Even allowing for many events that go undetected this satisfies the observed frequency requirements. The large scale heights of these objects above the galactic plane would also result in a roughly isotropic distribution for the strong events. At lower energies one might begin to see some association with the galactic disk depending on the actual scale heights involved. We would predict roughly 100 events per d having total integrated flux > 10^{-7} erg cm^{-2}.

What is the fate of neutron stars that accrete at rates differing markedly from $10^{-10} M_\odot$ yr^{-1}? While there may be interesting activity associated with semi-degenerate shell flashes in such stars, we do not envision ^{12}C detonation as occurring. For a very low accretion rate, $\dot{M} < 10^{-11} M_\odot$ yr^{-1}, the energy generation rate from carbon burning is so slow that a runaway cannot occur even when nuclear energy generation exceeds neutrino losses. Electron conductivity is very efficient in such cases and using the criterion of Giannone and Weigert[17] it seems that no runaway occurs. Instead, the carbon and oxygen are neutronised at $\rho \sim 2 \times 10^{10}$ g cm^{-3}. We note, however, that a very uncertain electron screening factor has a dominant role in this statement. For much larger accretion rates, $\dot{M} > 10^{-8} M_\odot$ yr^{-1} (for the assumed geometry), the temperature of the helium-burning shell approaches 10^9K (ref. 16) so that carbon can ignite before becoming extremely degenerate and a detonation will not ensue. In this regard, the accretion rate plays a similar role to that of the mass of the star in ordinary stellar evolution. It is well known that a massive star, $M \gtrsim 10 M_\odot$, will ignite carbon in its core non-degenerately, whereas in a less massive star, carbon will ignite explosively in highly degenerate conditions[21].

Giant X-ray pulses?

We note finally the possible relationship of the above model for γ-ray burst production to the recently observed giant X-ray pulses (see discussions in ref. 22). Although the observations are difficult to interpret at this early stage we feel that two categories of X-ray bursts are being observed: aperiodic (without periodicity) and quasi-periodic. We believe, in agreement with W. D. Evans, R. D. Brian and J. P. Conner (unpublished), that the aperiodic bursts are the soft energy tails of the spectra of distant γ-ray burst sources. In the preceding section it was suggested that a plausible number of such distant

outbursts for our Galaxy was ~ 100 d⁻¹. The distribution of these bursts should correlate with the galactic plane and might have a tendency to peak towards the galactic centre where the spatial density of neutron stars is expected to be higher.

The quasi-periodic sources pose a greater problem. Periodicities as short as ~ 15 s have been observed which rule out any model based on carbon detonation. In fact, unless the neutron star can somehow act as a storage battery, a simple calculation shows that bursts with periods of ≤ 30 s and durations of 1 s cannot be seen above the steady-state accretion luminosity for any nuclear powered model. This follows from the fact that gravitational accretion liberates 10–20% of the rest mass energy of the accreted material while all subsequent nuclear reactions generate < 0.7%. Thus even if hydrogen explodes and produces X rays with 100% efficiency the 15-s observations are very difficult to interpret in terms of a flash that burns only the material accreted during the interflash period. We believe, however, that these rapid pulses can result from a series of helium (and perhaps carbon) shell flashes occurring in semi-degenerate material with a total mass ~ 10²² g that has accreted over a period of time that is long compared with the interflash period. (Admittedly, it might be difficult to accumulate 10²² g without the nuclear burning having already occurred.) The neutron star stores the potential nuclear energy for a time during which its entire output is from the release of gravitational energy and then liberates it in a series of rapid flashes. More violent flashes may heat up the material to such an extent that the diffusion of energy outwards (and hence the return to an unstable configuration) takes longer than for weaker flashes. The fact that the time between bursts is longer for the more energetic bursts is an observed attribute of certain X-ray burst sources²². Clearly a large range and diversity of nuclear outbursts

of various energies and time scales are possible for the possible spread of accretion rates, neutron star masses, and magnetic field configurations. The subject is ripe for continuing experimental investigations and serious theoretical examination, and hopefully this somewhat speculative paper will encourage both. We especially urge that known γ-ray burst sites be continuously monitored for giant X-ray pulses and vice versa.

¹ Klebesadel, R. W, Strong, I. B, and Olson, R. A., *Astrophys. J. Lett.*, **182**, L85–L88 (1973).
² Strong, I. B, Klebesadel, R. W., and Evans, W D., in *7th Texas Symp Rel. Astrophys.* (edit. by Bergmann, P. G., Fenyves, E J., and Motz, L.), 145–159 (New York Academy of Sciences, 1975).
³ Cline, T. L , in *7th Texas Symp. Rel. Astrophys.*, 159–164.
⁴ Ruderman, M A., in *7th Texas Symp Rel Astrophys.*, 164–181.
⁵ Schmidt, W K H., Cline, T. L , Desai, V D., and Teegarden, B. J., *Bull Am. Phys Soc.*, *Series II*, 21, No. 4, 542 (1976).
⁶ Ramaty, R , and Cohen, J. M., in *Goddard Contributions to the Los Alamos Conference on Transient Cosmic Gamma and X-Ray Sources* (Goddard Pub. No. X-660-73-315, VI-V9, 1973)
⁷ Lamb, D. Q., Lamb, F. K , and Pines, D , *Nature phys Sci.*, **246**, 52–54 (1973)
⁸ Arnett, W. D , *Astrophys. Space Sci* , 5, 180–212 (1969).
⁹ Arnett W. D., *Astrophys. J* , **169**, 113–118 (1971)
¹⁰ Bruenn, S. W., *Astrophys J Suppl. No. 207*, 24, 283–318 (1972).
¹¹ Bruenn, S W., *Astrophys. J. Lett.*, 183, L125–L129 (1973).
¹² Buchler, J. R , Wheeler, J. C , and Barkat, Z , *Astrophys. J* , 167, 465–478 (1971).
¹³ Ivanova, I. N , Imshennik, V. S , and Chechotkin, V. M., *Astrophys. Space Sci* , 31, 497–514 (1974).
¹⁴ Shakura, N I., and Sunyaev, R A., *Astr Astrophys* , 24, 337–355 (1973)
¹⁵ Rosenbluth, M N , *et al* , *Astrophys. J.*, 184, 907–910 (1973)
¹⁶ Hansen, C J., and Van Horn, H M , *Astrophys J.*, 195, 735–741 (1975).
¹⁷ Giannone, P., and Weigert, A , Z *Astrophys.*, 67, 41–63 (1967).
¹⁸ Van den Heuvel, E. P. J , *Astrophys. J. Lett* , 198, L109–L112 (1975).
¹⁹ Blaauw, A., *Bull. astr. Insts Neth.*, 15, 265–290 (1961)
²⁰ Whelan, J., and Iben, I , *Astrophys. J.*, 184, 1007–1014 (1973)
²¹ Paczynski, B., *Acta astr.*, 20, 47–58 (1970)
²² *Physics Today*, 29 No 4, 17–20 (1976).
²³ Lewin, W. H G., *IAU Circ.*, No. 2922 (1976).

Tritium inventories of the world oceans and their implications

Robert L. Michel

Department of Chemistry, University of California, San Diego, California 92037

Tritium inventories are given for the Pacific Ocean from 51°N to 60°S. As expected, the highest inventories are found north of 20°N. Inventories drop rapidly from 20°N to 10°N. From 10°N southwards, inventories in the top 500 m are essentially constant. From these data, it is calculated that 107±21 kg of tritium were present in the Pacific Ocean in 1970. A total of ~300±80 kg of tritium was present on the Earth's surface in 1970. This indicates that 550±160 kg of tritium have been produced by nuclear detonations. It is estimated that 62% of the tritium in the Pacific Ocean was added through molecular exchange, 5% through runoff from the continents, and 33% as precipitation.

TRITIUM ($t_i = 12.3$ yr) is present on the Earth's surface mainly as HTO. The only significant natural sources are cosmic rays passing through the Earth's atmosphere and possibly accretion from the solar wind An equilibrium quantity of ~3.5 kg is present from these sources¹. Another source of tritium is the nuclear reactor. This source has some effect on the content of atmospheric hydrogen, but accounts for only minor amounts of HTO²,³. The major source of the tritium now present on the Earth has been the detonation of fusion bombs. The amount

of tritium produced by nuclear explosions is about two orders of magnitude higher than that produced by all the other sources. Estimates of the amount have varied from 800 kg (ref. 4) to 200 kg (ref. 5).

Most tritium has been deposited in the world's oceans by one of three processes: precipitation as rain or snow, runoff from the continents, or molecular exchange. Precipitation has been monitored in detail by the International Atomic Energy Agency and continental runoff is small in comparison. The amount of tritium added through molecular exchange is not well known⁶⁻¹⁰.

The La Jolla Tritium Laboratory has studied and used tritium as an oceanographic and geochemical tracer since the late 1950s. From 1968 to 1972, several series of depth profiles from various locations in the Pacific Ocean were analysed for tritium (Fig. 1). The oceanographic implications have been discussed elsewhere¹¹, but here we use the observations to study geochemical problems such as molecular exchange and tritium inventories.

Concentrations in top 500 m of ocean

Table 1 lists inventories of tritium expressed as numbers of atoms km⁻² in the top 500 m at various profile stations. The 500-m surface layer is used. Tritium below that depth, if present in at all measurable quantities, is brought in by lateral flow

Table 1 Tritium profile inventory in the top 500 m of the Pacific

Cruise	Year	Position	Inventory $(10^{19}$ atoms km$^{-2})$
Seven Tow	1970	50°52'N 171°31'W	23.4
Seven Tow	1970	44°59'N 167°08'W	32.0
OSN	1972–73	30°N 140°W	29.7
Seven Tow	1970	28°01'N 156°01'W	21.0
La Pared	1965	27°44'N 118°03'W	14.5
Seven Tow	1970	22°49'N 157°12'W	29.4
La Pared	1965	21°00'N 119°14'W	13.8
La Pared	1965	19°28'N 116°02'W	10.3
Seven Tow	1970	18°09'N 169°01'W	25.4
Scan	1969	15°22'N 169°33'W	23.1
Seven Tow	1970	13°47'N 166°40'W	15.6
South Tow	1972	10°18'N 123°10'W	2.4
Scan	1969	9°58'N 151°31'E	6.3
Scan	1969	9°30'N 166°31'E	4.6
Scan	1969	9°20'N 179°56'E	5.6
Seven Tow	1970	8°55'N 158°35'W	5.6
South Tow	1972	5°13'N 124°32'W	5.0
Seven Tow	1970	3°13'N 164°48'W	6.6
South Tow	1972	0°28'N 125°31'W	4.6
South Tow	1972	4°59'S 126°25'W	2.7
South Tow	1972	9°59'S 127°22'W	3.1
South Tow	1972	29°27'S 144°40'W	4.4
South Tow	1972	39°45'S 137°58'W	3.2
South Tow	1972	40°39'S 93°06'W	2.6
South Tow	1972	50°57'S 118°47'W	2.5
Eltanin	1972	54°07'S 172°55'E	3.9
Eltanin	1972	59°37'S 171°13'E	2.8

from polar latitudes and is not a result of surface input at the station.

Some profiles do not go as far as 500 m down. In some of these cases, reasonable estimates can be obtained if tritium concentrations decrease rapidly with depth. In no case should the error from these estimates exceed 10%. In some cases no estimates were possible and inventories are not given for those stations (see ref. 12 for the complete listing of data).

Correlation of data with latitude

Figure 2 shows a plot of tritium inventory against latitude. Tritium inventories are relatively constant from the subarctic Pacific to ~20°N, but further south they start to show a marked decrease. A minimum is reached at ~10°N, and throughout the Southern Hemisphere they are almost constant. Inventories are highest at northern latitudes because of the spring leak phenomenon[13], which results in a large input of tritium into the oceans from 30°N to 60°N. Oceanographic factors also affect the tritium distribution; tritium isolines and density isolines show strong correlations[11]. Density isolines start rising sharply south of 20°N and rise closest to the surface at 10°N. Tritium isolines show the same trend and a minimum in tritium inventories is found around 10°N.

Inventories for tritium below 500 m are significant only in the Southern Hemisphere. For the latitudes of 30°S to 50°S, inventories for below 500 m are approximately as large as for above 500 m (refs 12, 14). In the north, the tritium present below 500 m is negligible as a fraction of the total.

As these data cover most of the Pacific Ocean from North to South, it is possible to estimate the tritium burden of that ocean. All inventories are age corrected to 1970. The inventory is calculated assuming that the tritium inventory is constant to ±20% over a given latitude band. Only in the extreme east at 10°N does a large deviation occur. Carbon-14 values are also lower in the equatorial eastern Pacific[15] as a result of upwelling and influx of water from the South Pacific. There are two tritium profiles available from a GEOSECS station at 28.5°N (refs 16 and 17). These profiles have an average tritium inventory similar to profiles at equivalent latitudes presented here.

Michel and Suess have noted that surface tritium distributions have an east–west asymmetry with highest values in the east[11]. Also, caesium-137 determinations have shown that this

radioisotope is mixed deeper in the west and this dilutes its surface concentration[18,19], and, although vertical distributions vary, inventories seem not to be significantly different. The assumption of a constant inventory on an E–W line can be assumed to be correct within ~ ±20%.

At 28°N, one Seven Tow profile had a low inventory compared to other profiles near it. A profile taken near the same position 4 months later showed a completely different distribution, so the profile at 28°N has been deleted from calculations.

Total inventories

In Table 2, the inventories of tritium at various ranges of latitudes are given for the Pacific Ocean. The inventories include tritium below 500 m as well as in the top 500 m. It is assumed the tritium content of the deep water is negligible[16,17]. A total of 107±21 kg was found to be present in the Pacific in 1970, with > 80% of the tritium present in the North Pacific.

Table 2 Pacific Ocean tritium inventory (1970)

Latitude	Number of T atoms $(10^{19}$ km$^{-2})$	km$^2 \times 10^6$	Quantity of T (g)
< 30°N	32	30	48,000
30°N–20°N	30	14.6	21,000
20°N–10°N	15	16.6	12,500
10°N–0°	5	20.1	5,000
0°–30°S	4	49.8	10,000
30°S–50°S	5.8	26.6	7,700
< 50°S	3	19	1,900
			Total 107,000±21,000 g

Some idea of the total tritium present on the surface of the Earth can be obtained from these data and data of other researchers on tritium in other large bodies of water[14,20–23]. Most other work has been on the North Atlantic. The Atlantic north of 30°N has a consistently higher tritium inventory per square kilometre than the Pacific. South of 30°N inventories are similar. In calculating total inventories, the Atlantic is found to have an inventory 1.6 times as large as the Pacific, north of 30°N, and equal to the Pacific south of 30°N. Measurements are very sparse in the South Atlantic and inventories are assumed to be similar to those in the South Pacific at similar latitudes.

No tritium data are available from the Indian Ocean. It is assumed that inventories in the Indian Ocean are similar to those of the Pacific at similar latitudes. Most of the Indian Ocean surface area is south of the Equator where the tritium concentrations are low. Thus, the error in estimating the Indian Ocean inventory is small in terms of the total inventory. The Arctic Ocean inventory is calculated entirely on the basis of input by precipitation derived from the data of Schell and Sauzay[24]. Any input into the Arctic by molecular exchange should be small, as ice cover inhibits this process (R. Weiss, personal communication). Therefore it seems that the total tritium in the world's oceans is 250±50 kg (Table 3).

Continental tritium

A small fraction of the Earth's water supply is found in the active freshwater bodies of the continental land masses.

Table 3 World inventory of tritium (1970)

Pacific Ocean	107
Atlantic Ocean	120
Indian Ocean	20
Arctic Ocean	6
Worlds Oceans' Total	250±50 kg
Land water	45±25 kg
Air (HTO and HT)	3±11 kg
Total	300±80 kg
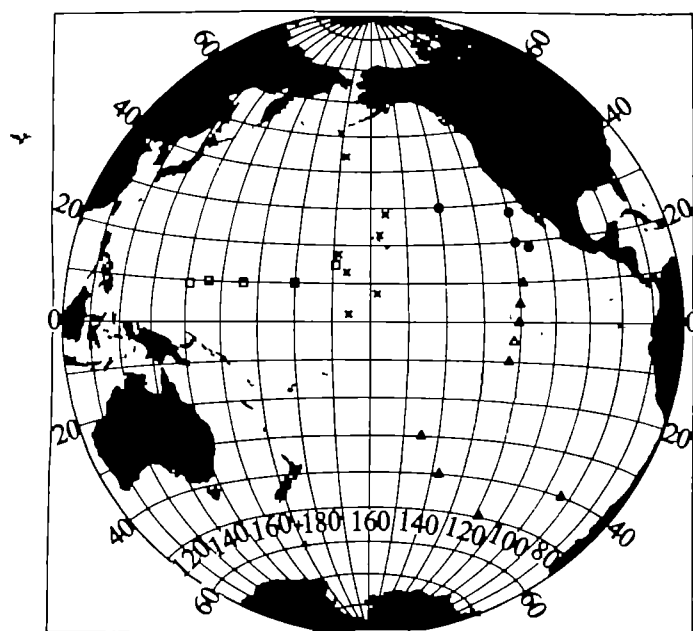
Fig. 1 Positions where tritium depth profiles were collected by the La Jolla Tritium Laboratory. ●, La Pared (1965); □, Scan (1969); ×, Seven Tow (1970); △, South Tow (1972); +, Eltanin (1972); O—OSN (1972).

Because of the continental effect on the tritium in rain, however, tritium concentrations are much higher in these water masses. Estimating the tritium inventory of the continental water masses is difficult, as concentrations there are much less uniform than in oceanic water masses. Also, very few data are available for 1970. It is possible to obtain an estimate of the tritium on the continents in the mid-1960s (ref. 25) which can then be age-corrected to 1970. Runoff and precipitation will cause some fluctuation in the inventory, but these effects will be small. It appears that there were 45±25 kg of tritium on the continents in 1970. The tritium content of air is known more accurately[2,3], and can be estimated to be 3±1 kg most of which is in the form HT.

Total tritium

These inventories give a total of 300±80 kg of tritium present on the surface of the Earth in 1970, most of it produced by nuclear testing.

From the world inventory, the total tritium produced by the early 1970s by nuclear explosions is calculated to be 550±160 kg,

Table 4 Modes of tritium input into the Pacific ocean (1970)

Precipitation	36 kg
Evaporation	−3 kg
Runoff	6 kg
Molecular Exchange (calculated)	68 kg
Total (from Table 2)	107 kg

where it is assumed that tritium production has been proportional to the megaton yield of nuclear explosives. Knowing the history of nuclear weapons testing, it is possible to correct for tritium decay[5].

This estimate may be low as it is assumed that no tritium is present in the deep water of the Pacific and Indian Oceans. Even a small concentration of tritium in the deep water would, however, result in a significant increase in the inventory. If the experimental upper limit of 0.2 TU (1 TU is one tritium atom per 10[18] hydrogen atoms) is used as the deep-water concentration, the figure for tritium produced by nuclear weapons testing would be 130 kg higher.

Molecular exchange

As mentioned previously, the extent of the contribution of molecular exchange to the input of tritium into the oceans is still uncertain. It was postulated[5,6] that most tritium in the oceans is added by molecular exchange. In a more formal treatment, Craig and Gordon[7] came to essentially the same conclusion. This seemed, however, not to be in agreement with experimental data. Two ocean studies[8,9] found experimentally that molecular exchange added < 0.33 of the tritium present. But the measurements were carried out in areas of low humidity where molecular exchange is expected to be low. Simpson[10] also found that molecular exchange could account for only about one third of the tritium added to Crater Lake. Conditions for this lake, however, are substantially different from those prevailing over the oceans, and the results are not necessarily significant for this problem.

Our data can be used to calculate the large scale molecular exchange for the Pacific Ocean. The only functions which affect tritium inventory are precipitation (P), evaporation (E), decay, runoff (R) and molecular exchange (ME). With the data corrected for decay, input by molecular exchange can be found by

$$ME = I - P - R + E \qquad (1)$$

where I is the inventory in 1970 (107 kg).

The amount of tritium deposited in the Pacific since 1963 can be calculated from the data of Schell and Sauzay[24]. Rainfall for most of the North Pacific is taken from data compiled by Reed and Elliott[26]. Rainfall data south of 20°N and evaporation data are taken from Sverdrup et al.[27]. Tritium concentrations at the sea surface are well known in the Pacific[12,13] so evaporation loss can be evaluated. The relative amounts of tritium added before 1967 have been calculated by Roether[11]. These data combined with the input data of Schell and Sauzay[24] have been used to estimate the input before 1963. All input by precipitation was corrected to establish the tritium deposited by 1970, and the net input was found to be 33 kg.

Approximately 9×10^{15} l of water enter the Pacific Ocean each year as runoff. Estimates of the tritium content of runoff are available from International Atomic Energy Agency (IAEA) data[28]. In calculating total runoff, it is assumed that runoff is of equal volume in all latitude bands from 50°N to 50°S. In

Fig. 2 Inventory of tritium in the top 500 m of the Pacific Ocean km[-2]. Symbols as for Fig. 1.
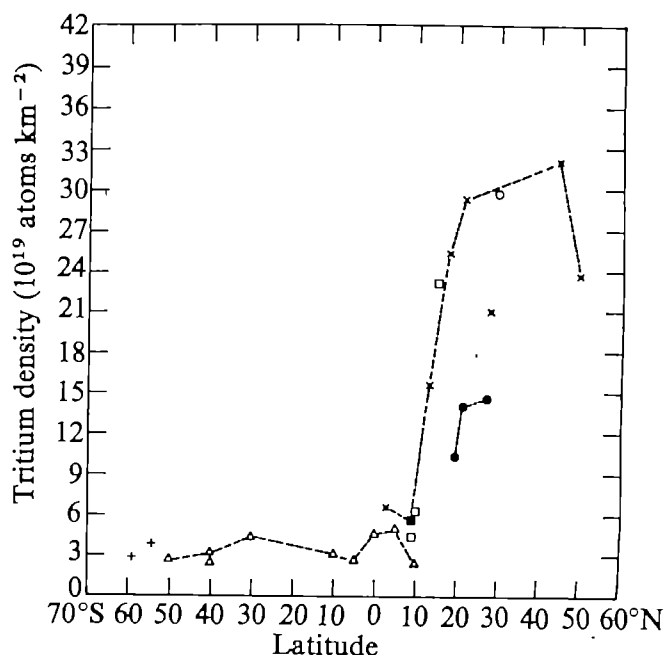
latitudes where no tritium data are available, it is expected that variations in tritium concentration in runoff are similar to the variations in rainfall. The errors are large but runoff tritium is small compared to input by molecular exchange or precipitation, being responsible for ~ 6 kg of the 1970 inventory. The input by molecular exchange is then found to be 68 kg (Table 4). This is 62% of the total tritium input in the Pacific—in agreement with previous estimates—and almost twice the input by precipitation alone. Thus, molecular exchange is indeed the predominant mechanism by which tritium enters the oceans.

This work was supported by the Chemical Oceanography Division of the NSF. Karen Dockins was in charge of technical operations of the La Jolla Tritium Laboratory until 1971, and then Teresa Broadwell took over. We thank H. E. Suess for discussions.

1 Lal, D., and Peters, B., *Encyclopedia of Physics*, 46/2, 551 (Springer-Verlag, Berlin, 1967).
2 Friedman, L., and Schulz, T. G., *J. geophys. Res*, 79, 785 (1974)
3 Ostlund, H. G., and Mason, A. S., *Tellus*, 26, 91 (1974).
4 Mrkal, J. A., in *Tritium*, 79 (Messenger Graphics, Las Vegas, 1973).
5 Erikson, E., *Tellus*, 17, 118 (1965).
6 Craig, H., *Phys. Rev.*, 105, 1125–1127 (1957).
7 Craig, H., and Gordon, L. L., *Rhode Island University, Narragansett Marine Laboratory Occasional Publications*, 3, 277 (1965).
8 Athavale, R. N., Lal, D., and Rama, *Proc. Ind Acad. Sci.*, A65, 73 (1967).
9 Gat, J. R., and Carmi, L., *J geophys. Res*, 75, 3039 (1970).
10 Simpson, H. J., *J. geophys. Res.*, 75, 5195 (1970).
11 Michel, R. L., and Suess, H. E., *J. geophys. Res.*, 80, 4139–4152 (1975).
12 Michel, R. L., thesis, University of California (1974).
13 Suess, H. E., *Science*, 163, 1405 (1969).
14 Tamuly, A., *Arctic*, 27, 27 (1974).
15 Linick, T. W., and Suess H. E., in *Proc. 8th Int. Conf. on Radiocarbon Dating*, I, C87 (Royal Society of New Zealand, Wellington, New Zealand, 1972).
16 Roother, W., Munnich, K. O., and Ostlund, H G., *J. geophys. Res.*, 75, 7672 (1970).
17 Roother, W., *Earth Planet. Sci. Lett.*, 23, 108 (1974)
18 Folsom, T., Mohanrao, G. J., Pillai, K. C., and Sreakumuran, C., *USAEC Health and Safety Laboratory, Fallout Program Quarterly Summary Report*, HASL-197 193 (1968).
19 Folsom, T., Sreekumaran, C., Hansen, N., Moore, J. M., and Grismore, R., *USAEC Rep., HASL-217* 1 (1970).
20 Ostlund, H. G., Dorsey, H. G, and Rooth, C. G., *Earth planet. Sci. Lett.*, 23, 69 (1974).
21 Roother, W, and Munnich, K. O., *Earth planet. Sci. Lett.*, 16, 127 (1972).
22 Bowen, V. T., and Roother, W., *J. geophys. Res.*, 78, 6277 (1973)
23 Ostlund, H. G, Broecher, R., and Peterson, W. H., *Tritium Laboratory Data Report* 3 (University of Miami, 1974).
24 Scholl, W. R., and Samay, G., *IAEA-SM-181/34*.
25 *Tritium concentrations in rain, rivers, oceans and other matter* (IAEA, published periodically since 1962).
26 Rood, R. K., and Elliott, W. P., *J. geophys. Res.*, 78, 7087 (1973).
27 Sverdrup, H. U., Johnson, M. W., and Fleming, R. H., *The Oceans, their physics, chemistry and general biology* (Prentice-Hall, Englewood Cliffs, New Jersey, 1942)
28 Dockins, K. O., Bainbridge, A. E, Hooterman, J. C., and Suess, H. E., *Tritium in the mixed layer of the North Pacific Ocean, Radioactive Dating and Methods of Low Level Counting*, 129–142 (International Atomic Energy Agency, Vienna, 1967).

Rolling hairpin model for replication of parvovirus and linear chromosomal DNA

Peter Tattersall*

Imperial Cancer Research Fund, Mill Hill Laboratories, Burtonhole Lane, London NW7 1AD, UK.

David C. Ward

Departments of Human Genetics and Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

A novel, quasicircular scheme is proposed for the replication of parvovirus DNA. Daughter strands are initiated after the copying and rearrangement of a terminal palindromic sequence, a process termed 'hairpin transfer'. Such a process may be involved in the replication of other viruses and host cell DNA.

ALL known DNA polymerases require a 3'-OH terminus on which to initiate strand elongation[1-3]. Although a satisfactory mechanism for the complete replication of circular DNA molecules can be envisaged by using RNA primers[3], additional molecular conjuring is required to explain the completion of the 5' termini of daughter strands in a linear DNA molecule. The problem arises because excision of a terminal RNA primer leaves a 5' terminal gap and because no polymerase is known which catalyses DNA chain growth in a 3' → 5' direction. Several elegant models have been put forward to circumvent the difficulty, all requiring special terminal structures, for example, terminal redundancy[4], palindromic termini[5] and covalently continuous (cross-linked) termini[6]. It has not been feasible to determine whether such terminal structures exist in the DNA molecules of eukaryotic cells, because of their extremely high molecular weight[7]. But practicable model systems for the study of chromosome structure and replication are provided by animal viruses with linear DNA genomes. For example, the parvoviruses contain linear single-stranded DNA genomes of only 1.4×10⁶–1.7×10⁶ daltons[8,9], which replicate

by way of linear double-stranded intermediates[10-18]. They are divided into two subgroups. The members of one subgroup, the adeno-associated viruses (AAV), are entirely dependent on adenovirus coinfection for their own replication, and package both plus and minus DNA strands in separate virions[8,9]. Members of the other subgroup, the autonomous parvoviruses, can replicate without a helper virus and package a unique viral DNA strand[8].

Structure of the parvovirus genome

To determine whether special arrangements of nucleotide sequence are necessary for the replication of a linear DNA molecule, we have examined the structure of the genome of the minute virus of mice[19] (MVM), an autonomous parvovirus. The results of this study[20] are summarised in Fig. 1. We have shown[20] that the 5' end of the molecule is a hairpin, represented diagrammatically by the sequence ABa. The duplex Aa can be isolated as an S1 nuclease-resistant, hydrogen-bonded fragment about 130 base pairs long, containing the 5' terminal base. The sequence represented by D is about 95% of the total genome. It is entirely single stranded and approximately 4,100 bases long. The 3' end can also form a hairpin denoted EFe, which serves as a primer for the synthesis of the complementary strand *in vitro* by a number of DNA polymerases. Various oncornavirus DNA polymerases (reverse transcriptases) can use MVM DNA efficiently as a primer-template. Since these polymerases lack exonuclease activity[21,22] and require base-paired primers[21-23], this implies that the 3' terminal base is hydrogen bonded in the duplex Ee. The length of this duplex is not precisely known, but it does not seem to be stable at

37 °C *in vitro*, as the entire genome, excluding the 5' hairpin, is sensitive to *E. coli* exo I (ref. 20). In addition, its priming ability is not destroyed after treatment with *E. coli* exo III.

Polymerases lacking 5' to 3' exonuclease activity synthesise a duplex DNA with a molecular weight 1.96 times that of the viral genome, in which the complementary strand is covalently attached to the template[29]. Restriction enzyme cleavage of this duplex has established that the viral genome is not circularly permuted[29]. The sizes of the single-stranded loops (B and F) are not known exactly, but they are less than 100 bases long, since they cannot be detected by electron microscopy. The observation that the loop F in the MVM duplex synthesised *in vitro* is extremely resistant to cleavage by S1 nuclease (D. C. W. unpublished results) suggests that this loop may be very small.

The structure determined for MVM DNA as shown in Fig. 1 has been demonstrated for several parvovirus genomes (M. B. Chow and D. C. W., to be published), and leads us to propose a model for the replication of such molecules. We have called this scheme the rolling hairpin as it is quasicircular and, in some respects, resembles the rolling circle model proposed for coliphage ΦX174 DNA replication by Gilbert and Dressler[34].

Rolling hairpin

The first step proposed for parvovirus DNA replication, shown in Fig. 2, is the gap-fill synthesis of the complement d of sequence D; as mentioned above, this step can be performed *in vitro* by several DNA polymerases[20]. This is followed by displacement synthesis and gap-fill synthesis, to copy the 5' terminal hairpin sequence aBA. To continue synthesis, we propose that this terminal duplex palindrome is now rearranged to form the 'rabbit-eared' structure shown in Fig. 2iv. This terminal synthesis and rearrangement recreates the hairpin originally present at the 5' end of the parental genome. At the
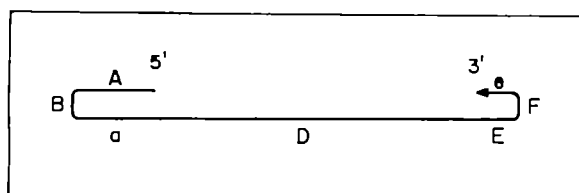
Fig. 1 Structure of MVM genome. In this and subsequent Figs, letters are used to represent unique blocks of sequence. The upper and lower cases denote complementary sequences, that is, if A is 5'-A-T-G-G-C-A-C-3' then a would be 5'-G-T-G-C-C-A-T-3'. The arrow denotes the 3'-OH end of a strand.

same time, it creates a copy of this hairpin, on the 3' end of the complementary strand, which can serve as a primer for the synthesis of the progeny strand. The new 3' hairpin eventually becomes the 5' hairpin of that future progeny genome, as shown in Fig. 2v. This process of 'hairpin transfer' is the essence of the model. The dimer length duplex is then completed by displacement synthesis. The resulting molecule comprises a single polynucleotide chain from which two viral genomes could be generated by suitable endonucleases; that is, the scheme shown in Fig. 2 represents one genome duplication. In the absence of endonuclease cleavage, however, this structure is obviously capable of continued synthesis by the same mechanism, yielding larger and larger multimers. In this respect our model strongly resembles the rolling circle model[34]. Thus a further roll around the 5' hairpin (that is, a repeat of steps *ii* to *vi* in Fig. 2) would generate the tetramer as shown in Fig. 3. In the concatemers formed by this replication process, the viral genomes are situated alternately on opposite strands of the duplex, with their palindromic ends overlapping.

If the single-stranded loop sequences B and F are themselves palindromes, with a phosphodiester bond as the axis of
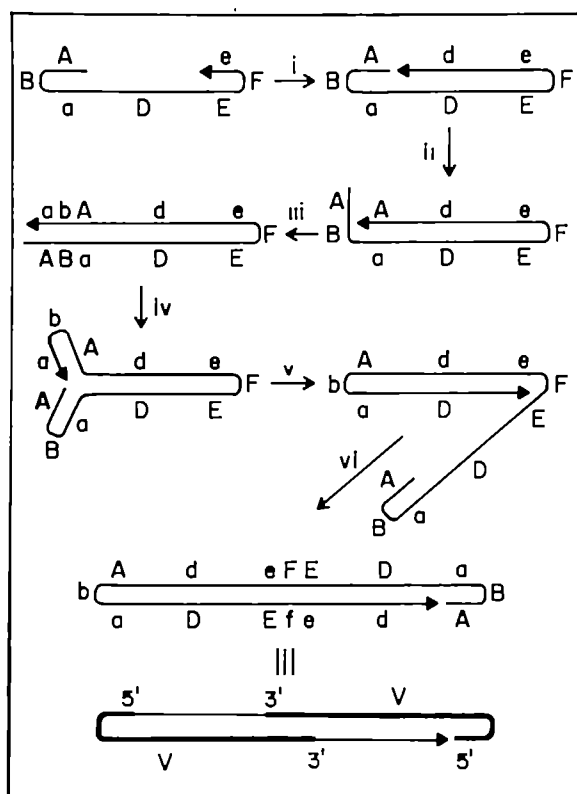
Fig. 2 The rolling hairpin model for parvovirus replicative DNA synthesis. *i*, Gap-fill synthesis; *ii*, displacement synthesis and *iii*, gap-fill synthesis to complete genomic hairpin; *iv*, enzymatic rearrangement of terminal palindrome to give rabbit-eared structure; *v*, displacement synthesis and *vi*, gap-fill synthesis The molecule is now essentially the same structure as the first intermediate between steps (*i*) and (*ii*), and repetition of steps (*ii*) and (*iii*) generates the dimer genomic hairpin. Repetition of steps (*iv*), (*v*) and (*vi*) generate the tetramer, and so on for larger multimers.

rotational symmetry (for example, GGTACC), there would be no difference between sequences B and b or between sequences F and f, and all genomes would be exactly the same; if they are not palindromes, then there would be four different kinds of viral sequence which differed in two small regions.

So far we have described a scheme for generating and replicating duplex forms of parvovirus DNA, based on the structure of the genome. The model now allows an explanation of some hitherto puzzling physical properties of parvovirus replicative intermediates. We showed previously[10] that some of the MVM duplex DNA molecules found in infected cells seemed to be concatemers up to about ten genome equivalents long. These molecules, however, contained no single strands longer than two genome equivalents, and on denaturation and reannealing formed predominantly monomer length duplexes. The long duplexes were also reduced to approximately monomer length by heating to submelting temperature. In addition we found that a large fraction of the MVM double-stranded DNA reannealed in a monomolecular fashion after denaturation, suggesting that viral and complementary strands were linked covalently, as predicted by the rolling hairpin model. Similar spontaneous renaturation behaviour has since been reported for duplex intermediates of parvovirus X 14 (ref. 14), AAV-2 (ref. 15) and Kilham rat virus[16]. Recently Straus *et al.*[16], have reported the isolation, from AAV-infected cells, of dimer length duplexes which renature spontaneously. These comprise a single DNA chain, four genomes long, containing alternating viral plus and minus strands. They have proposed a model for the replication of AAV DNA involving hairpin primers which is similar in parts to that proposed in Fig. 2. We originally suggested[10] that hairpin duplexes of genome length were formed by the use of the 3' end of the viral strand as a primer for the

synthesis of its complementary strand, and that this implied a mechanism for the regeneration of the primer in order to preserve the entire base sequence of the genome during replication. In the rolling hairpin model, this is accomplished by the process of hairpin transfer.

The overall physical properties of double-stranded DNA molecules replicating according to the model, such as the lengths of the single strands which the duplexes contain and the proportion of spontaneously renaturing molecules, will depend on how the intermediates are further processed to generate progeny genomes.

Synthesis and packaging of progeny viral DNA

A replication scheme which generates concatemers as replicative intermediates must necessarily include some process for excising progeny viral DNA. In parvovirus replication it is unlikely that this process relies only on the selection of genome size segments, since the viral DNA is not circularly permuted[20,26]. We envisage the first step in the process to be the introduction of a single-strand break, at the 5′ end of a genome within the

Fig. 3 Proposed mechanism for parvovirus progeny DNA synthesis, showing a tetramer intermediate. The progeny DNA strand, V, is indicated throughout as the heavy line. d represents the ~4,100 nucleotides of the complementary strand which is not packaged by the autonomous parvoviruses. The circle represents the capsid of an immature virion. *i*, Internal initiation of progeny strand displacement by sequence-specific nuclease cleavage; *ii*, displacement of progeny strand by elongation using the 3′-OH of the 'nick' as primer, and driven by concomitant packaging; *iii*, packaging of complete genome, and cleavage of 3′-OH end of genome (arrow); *iv*, release of intact virion and repair of the 3′ end of newly synthesised progeny strand to the 5′ end of its adjoining complementary strand; *v*, completion of a mature virion in the opposite direction and regeneration of the initial tetramer.

concatemer, that is between d and A in Fig. 3*i*. The 3′-OH terminus at this nick is then used as a primer for displacement synthesis of progeny DNA strands, giving rise to the duplex molecules with single-stranded side chains that are observed in infected cells[10,13,17]. In addition, we suggest that this displacement synthesis is driven by concomitant packaging of the progeny strand because little, if any, free single-stranded DNA can be detected during infection[13,14,17,18]. This idea is supported by the report[20] that a temperature-sensitive mutant of H-1 virus, with a virion which is physically unstable at high temperature, makes replicative intermediate, but not progeny DNA, at the non-permissive temperature.

After the 3′ terminal sequence e has been displaced, excision of the progeny genome could be completed by another sequence-specific nuclease, perhaps an integral part of the virion, releasing an intact virus particle and stopping the displacement synthesis. This second nuclease must be inactive on double-stranded DNA, because otherwise it would also introduce a nick in the template strand. The nascent 3′ OH terminal sequence e must be rapidly rejoined to the existing sequence d because the intact strand may be required as a template for the 3′ terminus of another genome being synthesised in the opposite direction, as in Fig. 3*iv*. If the packaging of progeny strands drives the displacement synthesis, the continued synthesis of complementary, minus strands (sequence d) is avoided, and the asymmetrical synthesis of progeny genomes, which has been reported for the autonomous parvoviruses[10,13,17] is assured. The specificity of such a packaging mechanism implies recognition of the 5′ end of the genome by the precursor virion. It is interesting that when infectious particles of MVM are lysed gently and examined in the electron microscope, a significant number of capsids remains associated with an end of the extruded DNA[30]. There is no detectable relationship between the two ends of MVM DNA[30], whereas the ends of most AAV genomes are related by an inverted terminal redundancy[27,28]. If terminal sequence recognition by the particle controls strand selection, this could explain why MVM only packages the plus strand, whereas AAV packages both plus and minus strands.

In concatemers where the 5′ ends of overlapping genomes have been nicked, displacement synthesis at one or both nicks will break the molecule in this region. Subsequent synthesis would regenerate the terminal palindrome of each of the fragments, which would then be able to continue replication after rearrangement as described in Fig. 2*iv*. Since none of the strands in double-stranded MVM DNA was found to be greater than dimer length[10], we suggest that the duplex concatemers had been nicked, perhaps during isolation, at the 5′ end of every component genome. Each continuous strand in such a molecule would have the sequence ABaDEFed and therefore comprise two palindromes. These sequences could rearrange, through cruciform intermediates, to produce monomer length hairpin duplexes at elevated temperature. This could explain the breakdown of MVM concatemers to monomers when they are heated to subdenaturing temperatures[10]. As both palindromic sequences in the concatemer would, in effect, become terminal hairpins in the monomer, one would expect this process to be essentially irreversible.

Hairpin transfer process

The central process of the rolling hairpin model is the synthesis and rearrangement of the palindromic terminus, to produce a 'rabbit-eared' structure. This permits the direction of chain growth to be reversed with respect to the original duplex. The rearrangement need not require complete melting out of the terminal region. If the loop sequence B is short it would be possible to interconvert the two structures through a cruciform intermediate, originally described by Platt[29] as twist transfer. Maniloff[30] has suggested, on theoretical grounds, that such a rearrangement should exhibit a transitional energy barrier and would be unlikely to occur spontaneously in physiological conditions. Therefore we have to postulate that there is an

Fig. 4 Proposed model for eukaryote chromosome terminal replication. *i*, Enzymatic rearrangement of terminal palindrome; *ii*, strand-displacement synthesis of daughter strand using 3'-OH of the parental strand as primer; *iii*, sequence-specific nuclease cleavage followed by *iv*, strand displacement and *v*, gap-fill synthesis to regenerate terminal palindrome.

enzyme activity which catalyses the rearrangement, and suggest that it is part of the DNA replication system of the cell which parvoviruses and possibly other viruses, require for rearrangement of their chromosomal ends during replication.

Hairpin transfer and host cell replication

To overcome the problem of replicating termini, originally discussed by Watson[4], Cavalier-Smith suggested[5] that the ends of linear chromosomes are palindromic so that, after excision of the 5'-terminal RNA primer, the exposed DNA strand could form a 3'-OH terminated hairpin. He proposed that this is then joined to the 5' end of the complementary strand by gap-fill synthesis and ligation. The palindromic sequence is then nicked at its 5' end by a sequence-specific nuclease. A combination of strand displacement and gap-fill synthesis, using the 3'-OH of the nick as primer, reconstitutes the parental strand.

Based on studies with MVM we have proposed a model for the replication of the parvovirus genome which overcomes the problem of the completion of linear termini by the process we have described as hairpin transfer. This concept can be extended to the replication of host chromosomes and other linear DNA molecules, and thus becomes a modification of Cavalier-Smith's model, involving fewer separate steps. In this scheme (Fig. 4), the steps of strand separation, RNA primer synthesis, excision, hairpin formation, repair synthesis and ligation, are replaced by the enzymatic rearrangement of the terminal palindrome. This produces a 'rabbit-eared' structure whose 3' end acts as a primer for the synthesis of the complementary strand. A sequence-specific nuclease then nicks at the 5' end of the palindrome, and synthesis using the 3'-OH of the nick as primer, reconstitutes the terminal sequence as before.

It is reasonable to expect that the expression of an enzyme system responsible for the initiation and completion of replication at chromosome termini would be coordinated with the passage of the cell through the mitotic cycle; in this respect it is important that the autonomous parvoviruses depend on, but cannot induce some host function(s) expressed transiently in the S phase[6]. The limited genetic capacity of these viruses suggests that they alter the host cell replication machinery very little. We suggest that the parvoviruses usurp the host's system for duplicating the termini of its chromosomes, to replicate their own DNA by the rolling hairpin process.

Terminal palindromes, which form terminal hairpins on strand separation, are emerging as a common feature of linear chromosomes. In addition to MVM DNA[20], such sequence arrangements have been demonstrated in the genomes of AAV[21] and herpesvirus[22]. The covalently continuous structure demonstrated for the vaccinia virus genome[23] could be regarded as a special form of terminal palindrome. These palindromes are found irrespective of the existence or polarity of terminally redundant regions in such molecules, either inverted as in adenovirus DNA[24,25], non-inverted as in herpesvirus DNA[22,34], or a mixture of both as found for populations of AAV DNA molecules[9,17,18,37].

Studies showing that a combination of exonuclease, polymerase and ligase can cross link the ends of coliphage T7 DNA[28,39] indicate that the T7 genome similarly contains a terminal palindrome, which agrees with the observation[40] that the termini of T7 DNA contain a self-complementary sequence. The isolation, from T7-infected cells, of spontaneously renaturing intermediates of greater than genome length[41] suggests that this sequence might be used for the replication of T7 termini also. Terminal palindromes may, therefore, be necessary for the replication of all linear chromosomes which do not, like phages λ or T4, replicate by becoming circular.

We thank Drs G. J. Bourguignon, L. Enquist and A. J. Shatkin for useful discussion and Drs J. Cairns and C. M. Radding for criticism of the manuscript.

1 Gefter, M. L., *A. Rev. Biochem.*, 44, 45–78 (1975)
2 Bollum, F. J., *Prog. Nucleic Acid Res. molec Biol.*, 15, 109–144 (1975).
3 Kornberg, A., *DNA Synthesis* (Freeman, San Francisco, 1974).
4 Watson, J. D., *Nature new Biol.*, 239, 197–201 (1972).
5 Cavalier-Smith, T., *Nature*, 250, 467–470 (1974).
6 Bateman, A. J., *Nature*, 253, 379–380 (1975).
7 Kavenoff, R., and Zimm, B. H., *Chromosoma*, 41, 1–27 (1973).
8 Rose, J. A., *Comprk. Virol.*, 3, 1–61 (Plenum, London 1974).
9 Berns, K. L., *Curr. Top. Microbiol. Immun.*, 65, 1–20 (1974).
10 Tattersall, P., Crawford, L. V., and Shatkin, A. J., *J. Virol*, 12, 1446–1456 (1973).
11 Saltzman, L. A., and White, W., *J. Virol*, 11, 299–305 (1973)
12 Siegl, G., *Pathol. Microbiol*, 40, 202–203 (1974)
13 Rhode, S. L., *J. Virol.*, 13, 400–410 (1974).
14 Mayor, H. D., and Jordan, L. E., *J. gen. Virol.*, 30, 337–346 (1976).
15 Sobring, E. D., Straus, S. E., Ginsberg, H. S., and Rose, J. A., *Fedn Proc.*, 34, 639 (1975).
16 Straus, S. E., Sobring, E. D., and Rose, J. A., *Proc. natn. Acad. Sci. U S A.*, 73, 742–746 (1976).
17 Siegl, G., and Gautschi, M., *J. Virol*, 17, 841–853 (1976).
18 Gunther, M., and May, P., *J. Virol.* (in the press).
19 Crawford, L. V., *Virology*, 29, 605–612 (1966).
20 Bourguignon, G. J., Tattersall, P J, and Ward, D. C., *J Virol.* (in the press)
21 Hurwitz, J., and Leis, J P, *J. Virol.*, 9, 116–129 (1972).
22 Baltimore, D., and Smoler, D., *Proc. natn. Acad. Sci. US A.*, 68, 1507–1511 (1971).
23 Goodman, N. C., and Spiegelman, S., *Proc. natn. Acad. Sci. U.S.A.*, 66, 2203–2206 (1971).
24 Gilbert, W., and Dressler, D., *Cold Spring Harb. Symp. quant. Biol*, 33, 473–484 (1968).
25 Berns, K. I., Kort, J, Fife, K. H., Grogan, E. W., and Spear, L., *J. Virol.*, 16, 712–719 (1975).
26 Rhode, S. L., *J. Virol*, 17, 659–667 (1976).
27 Kozmot, F. J., Carter, B. J., Garon, C F., and Rose, J. A., *Proc natn. Acad Sci. U.S A.*, 70, 215–219 (1973).
28 Berns, K I, and Kelly, T. J. Jr, *J. molec Biol.*, 82, 267–271 (1974)
29 Platt, J R., *Proc. natn. Acad Sci. U.S A.*, 41, 181–184 (1955).
30 Manlloff, J., *J. theor. Biol.*, 53, 497–501 (1975)
31 Denhardt, D T., Eisenberg, S., Bartok, K., and Carter, B, *J. Virol*, 18, 672–684 (1976)
32 Wadsworth, S, Hayward, G S., and Roizman, B., *J. Virol.*, 17, 503–512 (1976).
33 Geshhn, P., and Berns, K L, *J. molec. Biol.*, 88, 785–796 (1974).
34 Garon, C. F., Berry, K., and Rose, J. A, *Proc. natn Acad. Sci. U S A.*, 69, 2391–2395 (1972).
35 Wolfson, J., and Dressler, D, *Proc. natn. Acad. Sci. U S A*, 69, 3054–3057 (1972).
36 Grafstrom, R H., Alwine, J. C., Steinhart, W. L., Hill, C W, and Hyman, R. W, *Virology*, 67, 144–157 (1975).
37 Gerry, H. W., Kelly, T. J., Jr, and Berns, K. L., *J. molec. Biol.*, 79, 207–225 (1973).
38 Weiss, B., *Proc. natn. Acad. Sci. U.S.A.*, 65, 652–659 (1970)
39 Sadowski, P., McGeer, A, and Becker, A, *Can. J Biochem.*, 52, 525–535 (1974).
40 Obel, J., and Freifelder, D., *J. molec. Biol.*, 64, 439–447 (1972).
41 Barnhal, R., and Thomas, C. A. Jr, *J. molec. Biol.*, 51, 145–155 (1970).

Relationships between structure and activity of retinoids

Michael B. Sporn, Nancy M. Dunlop, Dianne L. Newton & William R. Henderson

Lung Cancer Branch, Carcinogenesis Program, National Cancer Institute, Bethesda, Maryland 20014

Structure–function relationships of 19 retinoids were studied in three in vitro systems in defined, serum-free medium. These systems measure effects on differentiation and growth, as well as toxic effects. The ring, side chain, and polar terminal group of natural retinoids all may be modified synthetically to yield highly active compounds. Several retinoid ethers were found to be significantly less toxic than the corresponding carboxylic acids or alcohols.

RETINOIDS (the set of molecules comprised of vitamin A and its synthetic analogues) are known to exert profound effects in controlling cellular differentiation in epithelium of many target tissues, including bronchi and trachea, uterus, prostate, kidney and bladder, stomach, intestine, pancreatic ducts, and skin[1]. It has recently been shown that several synthetic retinoids can prevent the development of epithelial cancer of the skin[2,3], respiratory tract[4], mammary gland[5,6], and urinary bladder[7] in experimental animals. In these experiments, the retinoids were not given to the animals until after completion of administration of the respective carcinogens. In these conditions, the retinoids do not prevent initiation of carcinogenesis, but rather, they modify preneoplastic states during the latent period of cancer development, before the onset of invasive malignancy. Although retinoids may act by more than one mechanism to modify preneoplastic states, their direct ability to control cell differentiation in target epithelia is undoubtedly of major importance in their suppression of carcinogenesis. For example, it has been shown that direct application of retinoids to prostatic[7,8] and tracheal[9] organ cultures can reverse hyperplastic or anaplastic lesions induced by chemical carcinogens; in these experiments, also, retinoids were not added to the cultures until after the lesions had been initiated by the carcinogen.

Many synthetic retinoids have now been made, in which there has been substantial modification of the ring, side chain, or polar terminal group of the natural vitamin A molecule. For both practical and theoretical considerations, an analysis of structure–activity relationships in this set of molecules is of fundamental importance. It is important to establish not only the molecular parameters that are required for retinoids to control epithelial cell differentiation and to prevent carcinogenesis, but also to establish the molecular parameters responsible for the undesirable toxic effects which retinoids may have on tissues, since such toxic effects may limit the effective use of retinoids in cancer prevention[10]. It is already clear that several synthetic retinoids, which have potent activity for the prevention of cancer in experimental animals, for the control of epithelial cell differentiation *in vitro*, or for reversal of premalignant lesions *in vitro*, have little capacity to promote growth in animals maintained on vitamin A-deficient diets[11]. Thus, classical nutritional methods (promotion of growth in vitamin A-deficient animals) may not be particularly useful to assess the potential activity of retinoids in cancer prevention, and new methods must be developed for correlation of structure and activity. Although measurement of ability to prevent various types of epithelial cancer in the intact experimental

animal is an important measure, such tests are very expensive and time-consuming when assaying large numbers of new compounds. Moreover, it is difficult to evaluate cellular and biochemical mechanisms of action in the live animal.

For these reasons, we have assayed the activity of retinoids in several organ and cell culture systems, using chemically defined, serum-free medium, in which structure–activity relationships can be evaluated quantitatively. Here we report on the biological activity of 19 retinoids in three such *in vitro* systems, which measure effects on both cell differentiation and growth (reversal of keratinisation in retinoid-deficient tracheal epithelium, increase of RNA in epidermal cell cultures), as well as the toxic effects on tissue (destruction of tracheal cartilage). More than 5,500 organ and cell cultures (including untreated control cultures) were used in these assays. Finally, since it is important that *in vitro* measurements be extended to evaluate utility *in vivo*, we also describe some major *in vivo* changes in tissue distribution of retinoids which result from synthetic modification of the polar terminal group.

Methods for *in vitro* studies

Structures of the retinoids (all provided by Hoffmann–La Roche Inc., Nutley, New Jersey and F. Hoffmann–La Roche & Co. A. G., Basel, Switzerland) used in the present study are shown in Fig. 1 and Table 1. Three series of retinoids have been studied: Fig. 1*a* retinoids which retain the β-cyclohexenyl ring and the all-*trans* side chain of the natural retinoids, but which have modifications of the polar terminal group; Fig. 1*b* retinoids in which a fluorine atom has been inserted into the all-*trans* side chain; and Fig. 1*c* retinoids which have a substituted
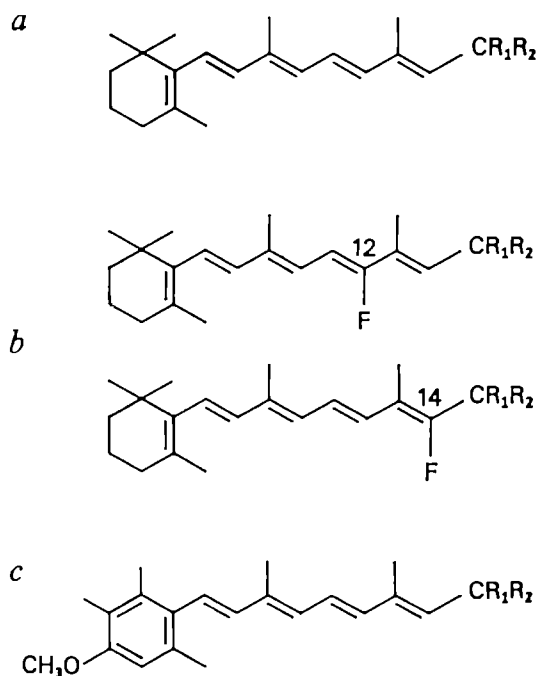
Table 1 Activity of retinoids in tracheal organ cultures and epidermal cell cultures

Structure, R1-, R2-	Trivial name	Reference	Reversal of keratinisation, tracheal organ culture, ED50(M)	Increase in RNA, epidermal cell culture, ED50(M)	Toxicity to cartilage, tracheal organ culture, TD50(M)
Fig. 1a, -H2, -OH	Retinol	26	2×10^{-9} (54)*	9×10^{-10} (64)	2×10^{-4} (67)
Fig. 1a, -H2, -OCH3	Retinyl methyl ether	20, 27	3×10^{-9} (122)	4×10^{-9} (83)	1×10^{-4} (103)
Fig. 1a, -H2, -OC4H9	Retinyl butyl ether	20	3×10^{-8} (55)	$>1\times10^{-6}$ (40)	not toxic at 1×10^{-4} (42)
Fig. 1a, -H2, -OCOCH3	Retinyl acetate	26	3×10^{-9} (123)	1×10^{-9} (18)	6×10^{-4} (63)
Fig. 1a, -H2, -NHCOCH3	N-Acetyl retinyl amine	28	9×10^{-9} (113)	$>1\times10^{-6}$ (52)	not toxic at 1×10^{-4} (50)
Fig. 1a, -H2, -N(CO)2C6H4	N-Retinyl phthalimide	28	5×10^{-7} (49)	inactive at 1×10^{-6} (48)	not toxic at 1×10^{-4} (10)
Fig. 1a, -H, =O	Retinal	26	3×10^{-10} (98)	8×10^{-10} (37)	5×10^{-7} (49)
Fig. 1a, -H, =NNHCOCH3	Retinal acetylhydrazone	29	4×10^{-10} (54)	3×10^{-9} (35)	7×10^{-7} (25)
Fig. 1a, -H, =NOH	Retinal oxime	30	1×10^{-8} (61)	9×10^{-9} (31)	2×10^{-6} (25)
Fig. 1a, =O, -OH	Retinoic acid	26	3×10^{-10} (337)	4×10^{-10} (287)	1×10^{-7} (116)
Fig. 1a, =O, -OC2H5	Retinoic acid ethyl ester	31	5×10^{-10} (119)	6×10^{-9} (100)	2×10^{-7} (18)
Fig. 1a, =O, -NHC2H5	Retinoic acid ethyl amide	32	2×10^{-9} (52)	$>1\times10^{-6}$ (70)	9×10^{-7} (32)
Fig. 1b, 12-fluoro =O, -OC2H5	12-Fluororetinoic acid ethyl ester	33	3×10^{-10} (63)	1×10^{-7} (40)	2×10^{-7} (18)
Fig. 1b, 14-fluoro =O, -OC2H5	14-Fluororetinoic acid ethyl ester	34	4×10^{-10} (68)	1×10^{-8} (40)	5×10^{-8} (18)
Fig. 1c, -H2, -OH	TMMP analogue of retinol	35	2×10^{-8} (48)	9×10^{-8} (39)	9×10^{-7} (34)
Fig. 1c, -H2, -OCH3	TMMP analogue of retinyl methyl ether	35	3×10^{-8} (70)	5×10^{-7} (38)	4×10^{-4} (57)
Fig. 1c, =O, -OH	TMMP analogue of retinoic acid	35	6×10^{-9} (187)	7×10^{-9} (86)	6×10^{-8} (88)
Fig. 1c, =O, -OC2H5	TMMP analogue of retinoic acid ethyl ester	35	2×10^{-8} (53)	2×10^{-7} (31)	2×10^{-7} (63)
Fig. 1c, =O, -NHC2H5	TMMP analogue of retinoic acid ethyl amide	35	3×10^{-8} (66)	5×10^{-8} (32)	1×10^{-4} (52)

ED50(M) for reversal of keratinisation in tracheal organ culture is the dose for reversal of keratinisation in epithelium of 50% of retinoid-deficient hamster tracheas using standard assay conditions[13]. ED50(M) for increase in RNA in epidermal cell culture is the dose for 50% increase in total RNA in mouse epidermal cells, also using standard assay conditions[14]. TD50(M) for toxicity to cartilage in tracheal organ culture is the dose for release of 20% of total proteoglycan from hamster tracheal cartilage into culture medium using standard assay conditions (see text).
*Numbers of organ or cell cultures used for evaluating each retinoid are shown in parentheses.

aromatic (trimethylmethoxyphenyl, TMMP) ring, as well as modifications of the polar terminal group. All retinoids were dissolved in dimethylsulphoxide (DMSO) before addition to cultures; an equivalent amount of DMSO (which never exceeded 1%) was added to all control cultures.

The test systems both for assay of reversal of keratinisation in organ cultures of tracheal epithelium from retinoid-deficient hamsters[12,13], and for assay of increase in RNA in cell cultures of mouse epidermis[14], have been described in detail. Typical dose–response curves for effects of retinoids in reversing keratinisation in tracheal organ culture are shown in Fig. 2; typical dose–response curves of their effects in increasing RNA in epidermal cell cultures have been published previously[13]. Toxicity of retinoids to cartilage of hamster trachea in organ culture was measured by a modification of a procedure for assay of retinoid toxicity to rabbit ear cartilage[11]. Briefly, tracheae from 4–5-week-old hamsters fed a normal laboratory stock diet were removed in sterile conditions, trimmed carefully of connective tissue, and then cultured for 3 d in the same medium used for the assay of keratinisation reversal, using air containing 5% CO2 as the gas phase. These tracheal organ

cultures were exposed to retinoids during the entire 3-d period Crystalline bovine serum albumin (0.5%) was added to some cultures to ensure solubility of retinoids at the highest concentrations used. After this period, both the trachea and the culture medium from each dish were saved; the amount of cartilage proteoglycan released into the culture medium, as well as that remaining in the cartilage, was measured with Alcian blue[15]. The total proteoglycan remaining in the cartilage was solubilised by papain digestion (digestions were performed at 37 °C for 2 h with shaking, using crystalline papain, 1 mg ml^{-1}, in 0.05 M sodium acetate buffer, pH 5.8, containing 0.001 M cysteine HCl). Chondroitin sulphate was used as the proteoglycan standard. At the end of the assay, the amount of proteoglycan found in the culture medium and the amount remaining in the cartilage were summed, and toxicity for any retinoid at any dose level was plotted (Fig. 3) as the percentage of the total proteoglycan released into the medium[11].

Reversal of keratinisation by retinoids

The ability of the respective retinoids to suppress keratinisation and reverse squamous metaplastic lesions caused by retinoid

Table 2 Comparison of tissue levels of retinoids 24 h after either a single or the last of nine oral doses

Retinoid	Single dose			Nine doses		
	Number of rats	Liver	Breast	Number of rats	Liver	Breast
Solvent control	18	0.37±0.02	0.0052±0.0003	19	0.53±0.02	0.0088±0.0015
Retinyl acetate	18	1.82±0.13	0.014±0.002	17	11.80±0.75	0.082±0.004
Retinyl methyl ether	17	0.92±0.10	0.060±0.01	10	7.38±0.98	0.71±0.10
Retinyl butyl ether	12	0.45±0.04	0.026±0.004	12	0.79±0.03	0.10±0.01

Sprague–Dawley female rats were dosed orally with retinoids (30 μmol each dose) dissolved in 0.4 ml of ethanol–trioctanoin (1:3). The rats receiving nine doses were treated three times weekly. All rats were killed 24 h after their final dose and tissue levels were determined with the trifluoroacetic acid method[37] after extraction of tissue with ethyl ether[36]. Retinyl acetate, retinyl methyl ether, and retinyl butyl ether all gave linear standard reactions with trifluoroacetic acid at 616 nm. Figures are total μmol of retinoid found per g wet weight tissue ± standard error.

deficiency in tracheal epithelium is shown in Table 1 and Fig. 2. This assay is a very sensitive measure of the ability of retinoids to control normal epithelial cell differentiation, since the same process can be shown to occur in the live animal[1]. Dose–response plots (Fig. 2) were made for all retinoids, and the values derived for the dose effective in suppressing keratinisation in half the cultures are shown in Table 1. Many retinoids were extremely active in the 10^{-9} M range (approximately 300 pg ml^{-1}).
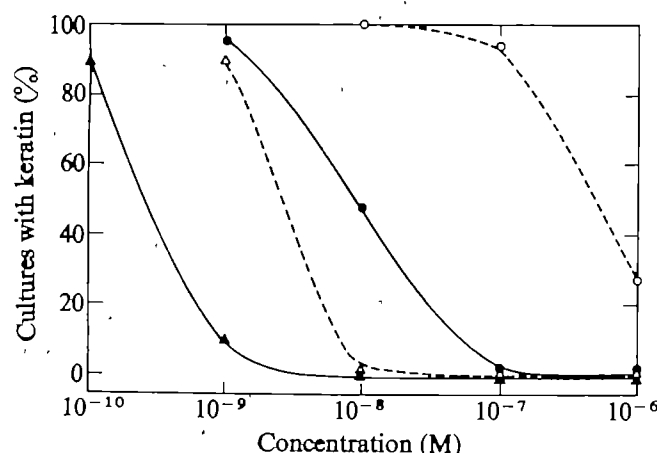
Fig. 2 Dose–response curves for reversal of keratinisation in organ cultures of retinoid-deficient tracheal epithelium by application of retinoids. Tracheae (one per culture dish) were treated with retinoids for 7 d before scoring for the presence of both keratin and keratohyaline granules. ▲, Retinoic acid; △, retinyl acetate; ●, N-acetyl retinyl amine; ○, N-retinyl phthalimide. Similar curves were obtained for all retinoids reported in Table 1, and the ED$_{50}$ was derived from these plots. Numbers of tracheae used in these experiments are shown in Table 1.

Several interesting results can be seen. In both series of retinoids—those with a β-cyclohexenyl ring (Fig. 1a), as well as those with an aromatic TMMP ring (Fig. 1c)—compounds with a carboxylic acid terminus were slightly more active than those with a carboxylic acid ethyl ester or carboxylic acid ethyl amide terminus, although the differences were not great. Also, in both series, compounds with terminal methyl ether functions were almost as active as the corresponding alcohols. Retinyl butyl ether was one-tenth as active as retinyl methyl ether or retinyl acetate. The oxygen atom on the terminal carbon of retinyl acetate may be replaced with an amine function, resulting in the compound N-acetyl retinyl amine, which was one-third as active as retinyl acetate. The N-phthalimide derivative was essentially inactive, except at the highest concentration tested. The acetylhydrazone derivative of retinal was almost as active as the parent aldehyde, although the oxime derivative of retinal had much less activity. The 12- and 14-fluoro-derivatives of retinoic acid ethyl ester were very active in reversing keratinisation.

RNA increase in epidermal cultures by retinoids

The ability of the respective retinoids to increase RNA in epidermal cell cultures is also shown in Table 1. There is a remarkable correspondence between activities as measured in the tracheal organ culture assay and the epidermal cell culture assay for 15 of the 19 retinoids. Three retinoids (retinyl butyl ether, N-acetyl retinyl amine, and retinoic acid ethyl amide) which reverse keratinisation in tracheal epithelium, were essentially inactive in the epidermal cell culture system, and a fourth compound, 12-fluororetinoic acid ethyl ester, was markedly less active in the epidermal cell assay. The reason for the above four exceptions is not known; but it is conceivable that epidermal cells from newborn mouse skin are deficient in the appropriate enzymes which might activate these four retinoids to biologically active forms in tracheal epithelium.

Toxicity of retinoids to tracheal cartilage

The toxicity of retinoids to tracheal cartilage *in vitro* is shown in Table 1 and Fig. 3. Dose–response plots (Fig. 3) were made for all retinoids, and the values derived for the dose that would release 20% of the total tracheal proteoglycan are shown in Table 1. Retinoic acid, retinoic acid ethyl ester, 12- and 14-fluororetinoic acid ethyl ester, the TMMP analogue of retinoic acid, and the TMMP analogue of retinoic acid ethyl ester were the six most toxic compounds. In both series of retinoids (Fig. 1a, c), the ethyl amide derivatives were markedly less toxic than their respective parent carboxylic acids. Similarly, the methyl ether derivatives were significantly less toxic than their respective parent alcohols, which in turn, were less toxic than the corresponding carboxylic acids. Two compounds, namely retinyl butyl ether and N-acetyl retinyl amine, which could control tracheal epithelial cell differentiation *in vitro* were virtually without toxicity to tracheal cartilage *in vitro*. These two compounds were also essentially inactive in the epidermal cell culture assay. All these data suggest that metabolic activation of retinyl butyl ether and N-acetyl retinyl amine may be required for their activity, and that, in some circumstances, there may even be a different set of metabolites which control cellular differentiation and cause cellular toxicity.

Pharmacokinetics of retinoids

The above *in vitro* studies clearly show a wide range of changes in activity resulting from modification of the structure of the retinoid molecule. Practical usefulness of retinoids for cancer prevention depends not only on a compound having the desired activity and lack of toxicity, but also on its pharmacokinetics *in vivo*. Thus, although retinyl acetate is highly active and relatively non-toxic (compared with retinoic acid) *in vitro*, it has the undesirable property *in vivo* of being stored in the liver, mostly as retinyl palmitate[16,17]. These elevated levels of retinyl palmitate can eventually cause serious toxic effects, both in the liver and systemically[17]. Moreover, because of the tendency of retinol and its esters to be stored in the liver, these retinoids may not reach specific target organs in doses as high as might be desired.

The greater effectiveness of retinyl methyl ether, compared with retinyl acetate, in prevention of mammary cancer in rats has recently been reported[6]. One possible explanation of these findings might be that greater effective levels of retinyl methyl ether than of retinyl acetate were attained in the mammary gland, so we carried out experiments to investigate this possibility. High doses of retinoids, corresponding to those used previously[6] when prevention of mammary cancer was investigated, were given in these experiments. After either single or multiple oral dosing of retinyl methyl ether, much greater levels of this compound (or one of its metabolites which react positively with trifluoroacetic acid, TFA) were indeed found in the mammary gland and adjacent fat pad, than after similar dosing with retinyl acetate (Table 2). Retinyl methyl ether *in vivo*, however, is known to be cleaved to retinol and then to be stored in the liver as retinyl ester[16,18]; results shown in Table 2 confirm this known ability of the liver to store large amounts of retinyl methyl ether or its metabolites. Thus, although retinyl methyl ether yields high retinoid levels in the mammary gland and adjacent fat pad, the storage of this compound and its metabolites in the liver suggests that, if retinyl methyl ether is fed for long periods, it may be as potentially toxic to the liver as retinol or retinyl esters.
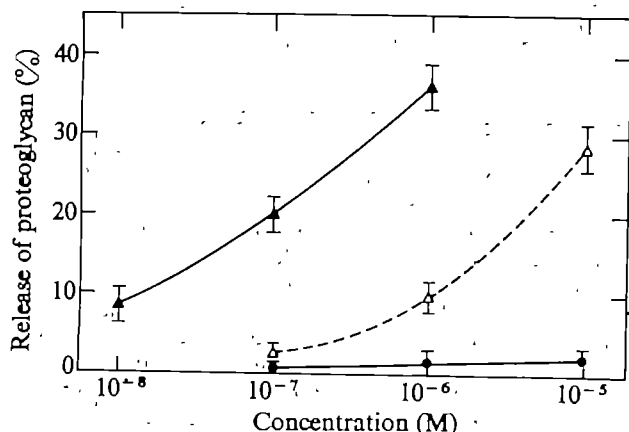
Fig. 3 Dose–response curves for toxic effects of retinoids to tracheal cartilage in organ culture. Tracheae (one per culture dish) were treated with retinoids for 3 d before measuring both proteoglycan released into the culture medium, and that remaining in the cartilage. ▲, Retinoic acid; △, retinyl acetate; ●, *N*-acetyl retinyl amine. s.e.m. of measurements is shown. Numbers of tracheae used in these experiments are shown in Table 1.

In contrast, another retinoid ether, retinyl butyl ether, seems to be metabolised by the intact animal in an entirely different way from retinyl methyl ether (Table 2). In spite of the fact that multiple dosing of retinyl butyl ether caused a twelvefold increase in the amount of TFA-positive retinoid in the breast, it did not even cause a twofold increase in liver retinoid in these experiments; in contrast, retinyl acetate and retinyl methyl ether respectively caused 22-fold and 14-fold increases in liver retinoid levels. The data indicate a major difference between the metabolism of retinyl butyl ether and of retinyl methyl ether, which we are now investigating using liver microsomes *in vitro*. Retinyl butyl ether thus seems to have the following unusual cluster of properties: it is a derivative of retinol with very low direct toxicity to tissue; it has significant growth-promoting activity in the whole animal[30]; and it does not accumulate in the liver after repeated oral dosing at high levels.

Conclusions

The data presented above show the complexity of relationships between structure and activity in retinoids. Significant modifications have been made in all three regions (ring, side chain,

and polar terminus) of the retinoid molecule and these modifications have resulted in biologically active molecules that may have practical use. Although it would not be justified to generalise to other retinoids from the present data, the following conclusions may be drawn from the present studies on retinoids with β-cyclohexenyl and TMMP rings: (1) the presence of a terminal carboxylic acid function conveys high activity, but is also associated with high toxicity; (2) fluorine atoms may be inserted into the side chain of β-cyclohexenyl retinoids to yield highly active compounds, of which the pharmacokinetics are at present largely unknown; (3) a terminal amine function, as in *N*-acetyl retinyl amine, is compatible with activity and may result in an analogue with low direct toxicity to tissue; and (4) modification of the polar terminal group can markedly lessen toxicity, and as shown with retinyl methyl ether and retinyl butyl ether, may also usefully affect pharmacokinetic properties.

The present data yield little information about the 'active form' of retinoids. A variety of retinoid structures are now known to be active in control of cell differentiation. The recent description and purification of several retinoid-binding proteins[31-36], which may have some resemblance to steroid-binding proteins, suggest that there need not be a single active form of retinoids. Rather, a wide range of retinoid structures may interact with cellular receptors directly in a biologically significant manner.

1 Moore, T., in *The Vitamins*, second ed, 1 (edit. by Sobrell, W. H., and Harris, R. S), 245–266 (Academic, New York, 1967).
2 Bollag, W., *Eur J. Cancer*, 10, 731–737 (1974).
3 Bollag, W., *Chemotherapy (Basel)*, 21, 236–247 (1975)
4 Port, C. D., Sporn, M. B , and Kaufman, D. G , *Proc. Am Ass. Cancer Res.*, 16, 21 (1975)
5 Moon, R C , Grubbs, C. J., and Sporn, M. B., *Cancer Res.*, 36, 2626–2630 (1976).
6 Grubbs, C. J , Moon, R. C., and Sporn, M. B., *Proc. Am Ass. Cancer Res.*, 17, 68 (1976).
7 Lasnitzki, I., *Br J. Cancer*, 9, 434–441 (1955).
8 Chopra, D. P., and Wilkoff, L J , *J. natn Cancer Inst.*, 56, 583–589 (1976).
9 Lane, B. P , and Miller, S L , *Proc Am. Ass Cancer Res* , 17, 211 (1976)
10 Sporn, M. B., Dunlop, N. M., Newton, D. L., and Smith, J. M , *Fedn Proc* , 35, 1332–1338 (1976).
11 Goodman, D. S., Smith, J. E., Hembry, R. M , and Dingle, J. T., *J Lipid Res* , 15, 406–414 (1974)
12 Clamon, G H., Sporn, M. B., Smith, J. M , and Saffiotti, U , *Nature*, 250, 64–66 (1974).
13 Sporn, M. B., *et al.*, *Nature*, 253, 47–50 (1975).
14 Sporn, M B , Dunlop, N. M., and Yuspa, S. H , *Science*, 182, 722–723 (1973).
15 Whiteman, P., *Biochem. J* , 131, 343–350 (1973).
16 Mahadevan, S., and Ganguly, J., *Biochem. J* , 81, 53–58 (1961).
17 Smith, F. R., and Goodman, D. S , *New Engl J. Med.*, 294, 805–808 (1976).
18 Thompson, J. N., and Pitt, G A J., *Biochim. biophys. Acta*, 78, 753–755 (1963).
19 Narindrasorasak, S., Pimpa, P , and Lakshmanan. M. R., *Biochem. J* , 122, 427–431 (1971).
20 Isler, O , *et al.*, *Helv. chim. Acta*, 32, 489–505 (1949)
21 Bashor, M. M., Toft, D. O , and Chytil, F., *Proc. natn. Acad. Sci. U.S.A.*, 70, 3483–3487 (1973).
22 Sani, B. P., and Hill, D. L., *Biochem. biophys. Res. Commun* , 61, 1276–1282 (1974).
23 Ong, D. E , and Chytil, F., *J. biol. Chem* , 250, 6113–6117 (1975).
24 Chytil, F., and Ong, D E , *Nature*, 260, 49–51 (1976)
25 Sani, B. P , and Hill, D. L., *Cancer Res* , 36, 409–413 (1976).
26 Schwartz, U., and Isler, O., in *The Vitamins*, second ed , 1 (edit by Sobrell, W. H., and Harris, R S), 5–101 (Academic, New York, 1967).
27 Hanze, A R., Conger, T. W , Wise, E. C , and Wembat, D I., *J Am. chem. Soc.*, 70, 1253–1256 (1948).
28 Kaenzle, F , *Helv. chim Acta*, 56, 1662–1671 (1973)
29 Miki, T , and Hara, Y., *Chem Pharmaceut. Bull.*, 10, 922–926 (1962)
30 Hubbard, R., *J Am. chem Soc.*, 78, 4662–4667 (1956).
31 Mayer, H., and Isler, O , in *Carotenoids* (edit. by Isler, O.), 325–575 (Birkhäuser, Basel, 1971).
32 *Chem Abstracts*, 75, 98697d (1971).
33 Machleidt, H., and Strehlke, G , *Justus Liebigs Annln Chem* , 681, 21–28 (1965)
34 Machleidt, H , and Wessendorf, R., *Justus Liebigs Annln Chem.*, 679, 20–26 (1964).
35 Belgian Patent No. 142589, March 29, 1974.
36 Ames, S R., Risley, H. A., and Harris, P L., *Analyt. Chem.*, 26, 1378–1381 (1954).
37 Dugan, R. E , Frigerio, N. A., and Siebert, J. M , *Analyt. Chem* , 36, 114–117 (1964).
38 Sporn, M.B., *et al.*, *Science* (in the press).

letters to nature

Do superheavy elements imply the existence of black holes?

THE discovery[1] of the superheavy elements 116, 124, and 126 raises the question of where these elements are likely to have been formed. The majority of the post-iron-peak nuclei are thought to have been produced in conditions of explosive nucleosynthesis (the r-process), particularly in conventional supernova explosions. The ability of the r-process to produce superheavy elements is, however, very uncertain[1]. The conditions necessary for superheavy element synthesis (β-decays occurring sufficiently slow that the $n \gamma \leftrightarrows \gamma n$ equilibrium is not disturbed) are difficult to realise in astrophysical situations. The n-process (J. B. Blake and D. N. Schramm, unpublished) requires less extreme conditions (the β decays are important) and may occur more often. The majority of the elements normally attributed to the r-process may have been synthesised in this way. Neutron-induced fission causes both processes to terminate at nuclei with high proton numbers, Z, but the n-process may allow it to reach the higher Z value.

We note that superheavy elements, such as those discovered and others much less stable in our environment, must exist in the outer layers of a neutron star[2]. Moreover, ideal conditions for the production of superheavy nuclei (high neutron flux and rapid β decays) are found in the disruption of a neutron star. The β decay rate is then fast compared with the expansion time scale (ref. 4 and J. M. Lattimer and D. N. Schramm, unpublished). We envisage such disruption being possible in either of two ways, both of which involve a black hole. Lattimer and Schramm (unpublished) and Lattimer *et al.*[4] have considered the tidal disruption of a neutron star from a close encounter with a black hole. Most of the disrupted star is swallowed by the hole, but some of the processed stellar material is assumed to escape.

Perhaps a more likely situation in which a neutron star is disrupted occurs when it accretes sufficient material that its mass exceeds the maximum mass for stable neutron stars. It has no alternative other than to collapse to form a black hole, and it again seems plausible that some of the outer layers are thrown off as it does so. The binary X-ray sources such as Her X-1 and Cen X-3 provide evidence that accreting neutron stars do exist.

The accretion process is enhanced if such binary systems tend to evolve towards coalescence, as is expected. One product of such evolution could be the giant stars envisaged by Thorne and Zytkow[4], in which accretion on to a neutron star core is a major power source for the whole star. Assuming that the neutron star cannot accept material at a rate much above that consistent with the Eddington limiting luminosity ($\sim 10^{38}$ erg s^{-1}), then $\sim 10^8$ yr are required before the collapse takes place. The resultant explosion could observationally closely resemble that of a more conventional supernova.

We thus argue that the most likely site for the production of superheavy elements is in the surface layers of a neutron star. The most plausible means by which these layers can be returned to the interstellar medium involve the intervention, or formation, of a black hole. We should, however, mention that in certain supernova models involving a hard equation of state, it is conceivable that the central regions collapse to neutron star densities, form superheavy elements, and then bounce. This may disrupt the original star entirely, leaving no remnant at all.

J. E. PRINGLE
D. S. P. DEARBORN
A. C. FABIAN

Institute of Astronomy,
Madingley Road,
Cambridge CB3 0HA, UK

1 Hodgson, P., *Nature*, 261, 627 (1976).
2 Schramm, D N., and Fisot, E O., *Astrophys. J* , 180, 551–570 (1973).
3 Baym, G., Pethick, C., and Sutherland, P , *Astrophys. J.*, 170, 299–317 (1971)
4 Lattimer, J M., Makle, F. D., Ravenhall, D. G., and Schramm, D. N , *Bull Am. astr. Soc.*, 7, 548 (1975).
5 Thorne, K S , and Zytkow, A. N , *Astrophys. J Lett* , 199, L19–L24 (1975).

Do superheavies come from neutron stars?

THE recent report of discovery of superheavy elements[1] (with nuclear charges near the predicted values for shell closure, $Z = 114$ and $Z = 126$), in surprisingly large quantities in terrestrial material, raises the question of the astrophysical sources for such nuclei. The very lightest nuclei may have been created in the first hours of the big bang, but significant amounts of elements beyond helium could not have been formed in these conditions. The heavier elements must have been formed in stars, and those as massive as uranium were probably formed by intense neutron bombardment immediately before or accompanying supernova explosions[1]. It seems unlikely, however, that superheavy nuclei could have been formed in this way, since attempts to reach the 'stability islands' by successive neutron captures must proceed by nuclei with very short lifetimes. This would almost certainly be the case for the island about $Z = 126$, if not for that at $Z = 114$. I would suggest that neutron stars might be a source for such superheavy nuclei.

Before theories of nucleosynthesis had reached their present state, Mayer and Teller considered evaporation from a fluid 'polyneutron' as a source of heavy elements[2], and came to the conclusion that nuclei with atomic masses of several hundred might be formed. With some qualification, such a polyneutron may be identified with a neutron star. We now know that the surface of a neutron star will very quickly be cooled below the crystalline melting point by neutrino emission processes[4], but it will be fluid for a short time after the collapse of a supernova core. Densities considerably higher than equilibrium values of perhaps $\rho \sim 10^{11}$ g cm^{-3} might be reached at the surface during oscillations following collapse. Evaporation of superheavies from the surface of such a newly formed neutron star would thus be virtually simultaneous with formation of heavy elements by neutron capture in the supernova explosion itself. Fission and β decay of the initial droplets of nuclear matter would be expected to yield some of the relatively stable elements about $Z = 114$, 126 and perhaps 164.

Another possibility is the disruption of a neutron star by tidal interaction with a black hole[5]. Matter can be ejected to infinity in such an encounter, and this matter might include

nuclear droplets large enough to give rise to the superheavy elements.

GEORGE L. MURPHY

Department of Physics,
The University of Western Australia,
Nedlands, Western Australia 6009

1 *Nature*, 261, 627 (1976); Gentry, R. V., *et al.*, *Phys. Rev Lett.*, 37, 11–15 (1976).
2 Trimble, V. T., *Rev. mod Phys*, 47, 877–976 (1975).
3 Mayer, M. G., and Teller, E., *Phys. Rev.*, 76, 1226–1231 (1949).
4 Dyson, F., *Neutron Stars and Pulsars*, Lecture 5. (Accademia Nazionale Dei Lincei, Rome, 1971)
5 Latimer, J. M., and Schramm, D. N., *Astrophys J. Lett.*, 192, L145–L147 (1974).

Spatial and temporal variations of the atmospheric sodium layer observed with a steerable laser radar

LASER radar observations of resonance scattering from atmospheric sodium during the night have, on occasions, shown evidence of short term changes in the height distribution[1–3]. It has been suggested that these changes arise from the movement of a horizontally structured sodium layer, but the measurements have all been limited to the vertical direction, and have provided no direct evidence of a horizontal variation. Evidence of horizontal structure at mesospheric heights has been provided by observations of hydroxyl emissions in the airglow, both from the ground[4] and from an aircraft[5,6]. In addition, the observed fluctuations of intensities and infrared rotational temperatures of the hydroxyl bands have been associated with the passage of atmospheric gravity waves[7,8]. We report here observations intended to examine the short term and small scale horizontal variations of the sodium layer which might also be associated with such waves.

The measurements were carried out using a laser radar system incorporating a flashlamp–pumped dye laser and a steerable mirror used for both transmission and reception. The details of the system will be described elsewhere. The steerable mirror was generally alternated between fixed angular positions with runs of ~ 500 laser pulses for each position. By this means, measurements of the height distribution of sodium at various horizontal separations were obtained.

The measurements made with the laser beam directed alternately towards the zenith and 10°E during the early morning of September 21, 1975 are shown in Fig. 1. For both directions the measurements taken during each 2–3-min interval are shown together with the standard errors; these results have been derived from the photon counts for each 1-km height interval, after allowing for the background noise

and correcting for the range dependence. The curves show, for comparison, the mean height distributions derived for each direction separately, the density scale being adjusted to give the best fit to the individual measurements near the peak. No attempt has been made to express the results in absolute sodium concentrations because of the possible variability of the atmospheric transmission coefficient and uncertainties arising from aerosol scattering in a calibration procedure based on returns from lower heights. The results shown for the two directions in Fig. 1 provide evidence of short term changes in the sodium layer which are different at locations separated by as little as 15 km at the height of the layer. Specifically, the sequence of measurements for the zenith, in comparison with the mean curve, indicates an upward movement of the layer between 0210 and 0214 or 0218 and a return to below the original height by 0227. The corresponding change in the data for the 10°E direction is seen as a downward movement between 0212 and 0216; a return to the original height, and even an upward movement, is suggested in the data for 0231. These results imply vertical movements of ~ 2 km in the sodium layer, these movements being of opposite sense at points separated by ~ 15 km. Furthermore, the suggestion of a reversal in the sense of the movement near the end of the observing period for each direction is consistent with an oscillation of period ~ 20 min. This is the first time such a small scale horizontal structure has been detected in the atmospheric sodium layer. With the data available it is not possible to distinguish between a horizontally moving structure and the passage of an atmospheric gravity wave. We note, however, that a simple interpretation of such a periodic oscillation with a horizontal wavelength of ~ 30 km in terms of a gravity wave propagating in the absence of a horizontal wind would require a vertical wavelength of ~ 7 km, whereas the vertical motions seem to be cophasal over the whole weight range of the sodium layer.

Our programme of measurements, which were carried out over eight nights during the period September 1975–January 1976, showed that, in general, the sodium layer showed little small scale spatial variation or temporal changes over periods up to ~ 1 h. This is illustrated in Fig. 2 which shows the measurements carried out on October 5, 1975 with the laser beam directed 10°E and 10°W of the zenith, corresponding to a horizontal separation of ~ 30 km at the height of the layer. The results for each direction show no evidence of any pronounced variation of the layer over the 40-min observing period. Moreover, a detailed comparison of the data on which the mean curves for the two directions are based shows that the differences are not significant. On two of the nights of observation, beam inclinations of 30° to the zenith were used, corresponding to a horizontal separation of ~ 100 km at the height of the layer. Even with this separation, significant differences
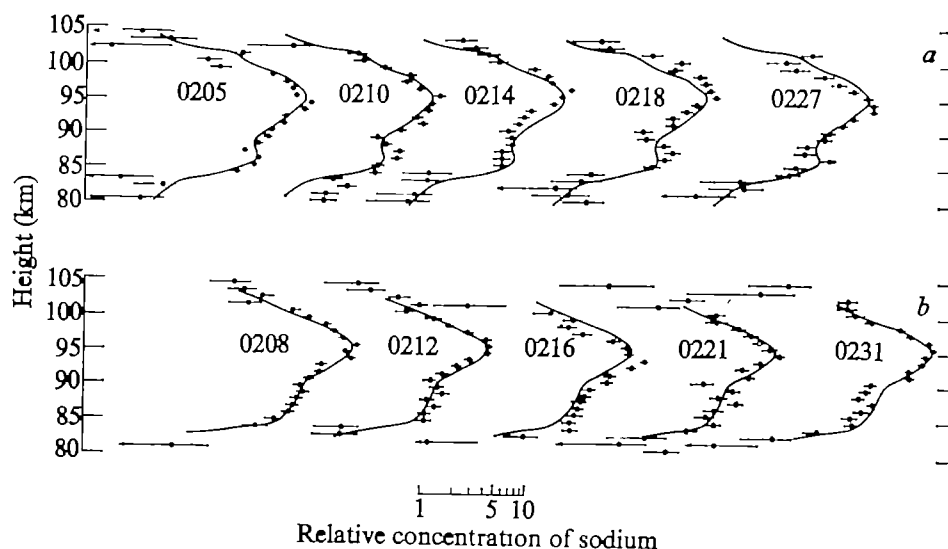
Fig. 1 Variations of the height distribution of atmospheric sodium in the zenith (*a*) and at 10°E to the vertical (*b*), corresponding to a horizontal separation of ~ 15 km at a height of 85 km. The curves represent the mean of all observations shown at each angle. Data for September 21, 1975.
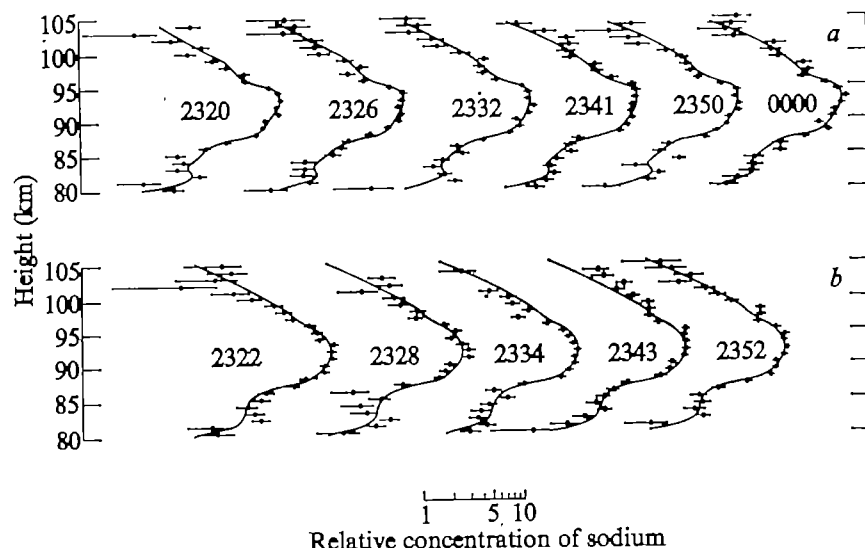
Fig. 2 Variations of the height distribution of atmospheric sodium at 10°E (*a*) and 10°W (*b*) to the vertical corresponding to a horizontal separation of ~ 30 km at a height of 85 km. The curves represent the mean of all observations shown at each angle. Data for October 5, 1975.

were observed only at heights < 82 km. In connection with this lack of variability, it should be noted that previous measurements at middle latitudes of the height variation of atmospheric temperature[9] and of very low frequency waves reflected from the D region[10] have indicated that the greatest degree of variability at mesospheric heights would be expected during winter months. It is intended to extend this type of measurement with improved system sensitivity and greater horizontal separation of the region being sounded.

We acknowledge the contributions of M. C. W. Sandford and E. Hammond to the design and construction of the laser radar system.

L. THOMAS
A. J. GIBSON
S. K. BHATTACHARYYA

Appleton Laboratory,
Ditton Park, Slough,
Berkshire SL3 9JX, UK

1 Sandford, M. C. W., and Gibson, A. J., *J. atmos. terr. Phys.*, **32**, 1423–1430 (1970).
2 Hake, R. D., Jr, *et al.*, *J. geophys. Res.*, **77**, 6839–6848 (1972).
3 Blamont, J. E., Chanin, M. L., and Megie, G., *Annls Geophys.*, **28**, 833–838 (1972).
4 Peterson, A. W., and Kieffaber, L. M., *Nature*, **242**, 321 (1973).
5 Peterson, A. W., and Kieffaber, L. M., *Nature*, **257**, 649–650 (1975).
6 Crawford, J., Rothwell, P., and Wells, M., *Nature*, **257**, 650–651 (1975).
7 Krassovsky, V. I., *Annls Geophys.*, **28**, 739–746 (1972).
8 Krassovsky, V. I., and Shagaev, M. V., *Planet. Space Sci.*, **22**, 1334–1337 (1974).
9 Theon, J. S., Nordberg, W., and Katchen, L. B., *J. atmos. Sci.*, **24**, 428–438 (1967).
10 Budden, K. G., Ratcliffe, J. A., and Wilkes, M. V., *Proc. R. Soc.*, **171**, 188–214 (1939).

Atmospheric temperature calculated for ozone depletions

THE proposal that the Earth's protective ozone layer is being depleted by free chlorine produced by the photolytic decomposition of chlorofluoromethanes (CFCl₃ and CF₂Cl₂, R11 and R12) has been the subject of much discussion and controversy[1–3]. Although the effects of ozone removal on human health is a prime concern, attention has also been directed towards the possible influence on the stratospheric temperature profile[4] with associated consequences for the climate. In previous papers[4,5] it was shown that uniform depletion of stratospheric ozone leads to cooling of the stratosphere. To obtain a realistic bound on the maximum extent of stratospheric cooling, we have computed the thermal effects resulting from model D of Wofsy *et al.*[1] for ozone depletion caused by chlorofluoromethane decomposition.

In model D the authors calculated the time dependence of the atmospheric ozone profile for an assumed 10% yr⁻¹ growth rate of R11 and R12 usage (equivalent to doubling every 7 yr), assuming the 1972 emissions were 2.2×10^5 and 3.5×10^5 tonnes respectively. The decrease from the 1972 ozone profiles are shown in Fig. 1 for the years 1979, 1984, and 1996 (the ozone profiles were obtained from S. C. Wofsy, private communication).

As an input to their calculations, Wofsy *et al.*[1] used a temperature profile which was based on the 1966 standard atmosphere and assumed to remain unchanged through 1996. They report that their model gives a good representation of the ozone profile for the undisturbed atmosphere, when compared with available data[7].

To establish internal consistency in our results, we first computed a 1972 temperature profile using the Manabe–Wetherald (M–W) model[6], and the Wofsy *et al.* 1972 unperturbed ozone profile. The results were in essential agreement with Wofsy's assumed temperature profile. Next the 1979, 1984 and 1996 temperature profiles were computed assuming only that the ozone profiles changed in accordance with Wofsy's model D prediction (Fig. 1). The results are shown in Fig. 2.

As shown before[4], an appreciable temperature decrease (of the order of several degrees) was calculated in the stratosphere while a slight but uniform temperature increase was calculated throughout the troposphere. The largest change in magnitude was ~ −7.5 °C at 40 km, for a 33% ozone depletion (1996). This is similar to the results of McElroy *et al.*[7] for the stratospheric temperature effects of

Fig. 1 Percentage of ozone decrease[4], based on R11 and R12 increasing at a rate of 10% yr⁻¹ (model D).
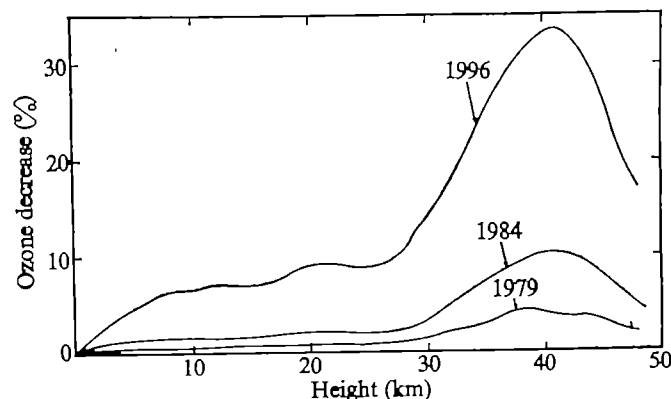
SST fleet operations, who found a maximum temperature decrease of ~ −10 °C at 40 km. In both cases the stratospheric cooling results from a decrease in the absorption of visible and infrared radiation by ozone.

The increased uncertainty in our calculated temperature change (Fig. 2) for 44 and 48 km reflects the difficulty in treating adequately the increasingly complex spectra of the optically active atmospheric constituents at high altitudes.

Our calculated tropospheric temperature changes were 0.04 °C for 1979, 0.07 °C for 1984 and 0.16 °C for 1996, a roughly linear increase with time, at a rate of ~ 0.06 °C/10 yr. This slight surface heating is the result of more solar radiation being absorbed by the particulate layer which, in turn, heats and radiates more infrared radiation downwards to raise the surface temperature.

Similar calculations for the ozone profiles of ref. 8 give a surface temperature increase of 0.14 °C for his case I and 0.15 °C for his case II. In case I, Crutzen assumed a fixed yearly Freon release rate (equal to the world 1972 production rate) into the troposphere and determined that a steady state would be reached by the year AD 2055. Halocarbons would then be diffusing out of the troposphere into the stratosphere at the same rate at which it is being released

Fig. 2 Calculated temperature changes, based on the calculated ozone profiles of Wofsy, McElroy and Sze[5]. The 1972 temperature profile was calculated using the CO_2 and H_2O profiles described previously[4], three layers of water clouds, a surface albedo of 0.097, a background low-lying particulate layer of optical density 0.065, and 4% Rayleigh scattering. The zenith angle and length of daylight were characteristic of 30°N latitude, July. Although the M-W model does not include general circulation, it does introduce a convective contribution when the lapse rate exceeds the thermodynamic adiabatic limit.

at the surface. In these steady-state conditions he calculated that the ozone would be reduced by 7.0% of its 1972 level. For production increasing at 10% yr⁻¹ (Wofsy, model D) this change would be reached by 1987.

The Crutzen ozone profiles were also used to test the sensitivity of the temperature calculation to surface albedo. As the surface albedo was increased from 0.1 to 0.6, the stratospheric cooling increased ~ 0.4 °C while the tropospheric heating decreased ~ 0.06 °C. Thus the dependence on surface albedo seems to be of relatively minor importance in the stratosphere and of greater importance for surface temperatures.

These calculations suggest that a drop of several degrees in the stratospheric temperature might occur by the end of the century if ozone depletion proceeds according to model D of Wofsy *et al*. More recent reaction rate constants are felt to reduce these ozone depletion rates by a factor of ~ 2 (see ref. 9) and it has been found[10] that the photocross-sections for R11 and R12 are reduced at lower temperatures. In addition, it has been suggested that $ClONO_2$ may be of significance in the photochemistry of stratospheric O_3. These effects, of course, were not included in the Wofsy's calculated ozone reductions, which I have used. The importance of this temperature effect on climate (if any) has not yet been established. Since previous calculations have indicated that the thermal effect may be strongly influenced by the altitude of the maximum in the ozone profile[4], the results obtained here would be expected to vary with both latitude and season.

RUTH A. RECK

Physics Department, Research Laboratories,
General Motor Corporation,
Warren, Michigan 48090

1 Rowland, F. S , and Mohns, M. J., *Rev. Geophys. Space Phys*, 13, 1–35 (1975).
2 Cicerone, R. J., Stolarski, R. S., and Walters, S., *Science*, 185, 1165–1167 (1974).
3 *Findings on the Effects of Chlorofluoromethanes on Stratospheric Ozone* (NAS, to be available, August 1976).
4 Reck, R. A., *Science*, 192, 557–559 (1976).
5 Wofsy, S. C., McElroy, M. B , and Sze, N. D., *Science*, 187, 535–537 (1975).
6 Manabe, S., and Wetherald, R T., *J atmos. Sci.*, 24, 241–259 (1967).
7 McElroy, M. B , Wofsy, S. C., Penner, J. E., and McConnell, J. C , *J atmos. Sci.*, 31, 287–303 (1974).
8 Crutzen, P. J., *Geophys Res. Lett.*, 1–4, 205 (1974).
9 *Fluorocarbons and the Environment*, 26–9 (Report of Federal Task Force on Inadvertent Modification of the Stratosphere, June 1975).
10 Looper, E. M., *Bioscience*, 25, 753–755 (1975).

A possible Himalayan microcontinent

ONE of the most significant features of the Himalayan mountain chain is its crystalline core which separates the tectonic belts of the Palaeo–Mesozoic–Cainozoic rock sequences of the Himalaya in the south and those of the Trans-Himalaya in the north. This crystalline core is described as the Central Crystalline Zone[1] and the Central Barrier[2]. There are indications that the rocks of this zone are essentially a Precambrian–Proterozoic rock suite[1,2], although they were strongly remobilised during the Himalayan orogeny. No hypothesis on the evolution of this tectogene has yet explained satisfactorily the existence of the crystalline core[4,5], and taking this and the tectono-stratigraphic evidence into consideration and identifying some constraints in the existing models on the origin of the Himalaya, an alternative model, based on plate tectonics, involving microcontinents, is suggested here.

In the geosynclinal models the Central Crystalline Zone has been given distinct geotectonic status as the Central Himalayan geanticline[6], intrageanticline[7], and miogeanticline[8], but its role in such models is uncertain. There have been attempts also to fit the stratigraphic and tectonic picture of the Himalaya into an Alpine framework[9,10]. Such an attempt runs into difficulties because, although the Himalaya is a Cainozoic mountain chain, the extent and type of Meso-Cainozoic sedimentary and tectonic record do not permit a satisfactory application of the classic Alpine geosynclinal model in the Himalaya[11]. The bulk of the rocks in the Himalaya *sensu stricto* are reworked Precambrian/Proterozoic basement rocks, and Cainozoic granite–granodiorite massifs, which occur in the Axial and Inner Tectonic Belts[12], roughly defining the core of this chain (Fig. 1). A narrow zone of tectonised Palaeo–Mesozoic rocks, having a marked lateral inconsistency and strong stratigraphic breaks at places, occur to the south of these belts as allochthonous masses, thrust over the Cainozoic sediments of the Foothills. The Palaeo–Mesozoic sedimentary record is, however, fairly complete in the Tibetan zone of the

Trans-Himalaya. The most prominent and continuous dislocation zones are the Main Boundary Fault (tectonic contact of the allochthonous Pre-Tertiary rocks with the Tertiary rocks), Main Central Thrust (tectonic contact of the Axial gneisses and granites with the reconstituted Precambrian–Proterozoic rocks of the Inner Belt), Trans-Axial Thrust (tectonic contact of the Tethyan sedimentary sequence with the Axial rocks). North of the main Himalaya, the Indus–Tsangpo suture[1,18] possibly defines an ophiolite-melange zone. These dislocation zones and the character of the rock suites involved between them, seem incongruous in the regional perspective of eu–miogeosynclinal couple which is one of the fundamental tenets of the geosynclinal hypothesis.

In the plate tectonic model, the Himalaya is considered to be the classic example of a continent–continent collision system[14,15] although such a system is not yet fully understood[16]. The application of such a model[17] is faced with no fewer difficulties than those in the geosynclinal models. The plate tectonic models postulate the existence of a Tethys ocean, thousands of kilometres wide, separating India and Eurasia, before the foundering of the Gondwanaland[13,18,19], and therefore, the Himalaya is considered to have originated at the site of this consumed ocean. The crystalline axis of the Himalaya and the Palaeo–Mesozoic shallow marine and even continental sedimentary sequence to the South and North of this axis impart the main geological constraint to such models. Moreover, Tibet is mildly seismic[20] and no Tertiary to Sub-Recent volcanics are recorded in this region. The gravity data in the Himalaya clearly isolate it geophysically from the Trans-Himalayan region[21]. These facts do not fit well in a scheme involving direct collision between the Indian and Tibetan shields. Plate tectonic models are in a state of confusion, which can be realised from the fact that it is still a matter of conjecture and personal conviction as to where the junction between the Indian and Eurasian plates actually lies. The suggestions are that it lies along the Indus–Tsangpo suture[13], the Main Central Thrust[22] and even the Tien Shan[23]. These difficulties have apparently led to the suggestion[24] of a Greater India in the Gondwanaland, and to the proposal[25] that the Himalaya might represent an intracontinental mountain chain. Because of this confusion, and

considering the stratigraphic, palaeontological and palaeomagnetic data, the idea of wide oceanic separation between India and Tibet has also been criticised[15,28].

It seems that many of these inconsistencies and difficulties in the concept of evolution of the Himalaya can be overcome if this tectogene is considered to have originated through the interaction of the Indian shield and a cluster of microcontinents, instead of a direct collision between the Indian and Tibetan shields, across the Tethys. Such an interaction would account for, among other geological and geophysical attributes which characterise the Himalaya and isolate it from other segments of the Mediterranean mobile belt, the crystalline core with Cainozoic granite–granodiorite intrusives, shallow marine and platform sediments on either side of this core. The presence of a microcontinent would also imply a reduction of effective width of the Tethys which is necessary[13,28] to explain the geological picture of this region.

A strong tectonometamorphic break[27] in the Himalayan stratigraphy is indicated by extensive Permo–Carboniferous conglomerate and boulder-bed (Blaini–Rangit–Lachi). This break is linked with the Hercynian hiatus after which Gondwana sedimentation had commenced in Peninsular India. A correlation between the Precambrian–Proterozoic and Palaeozoic rocks in the Himalaya and those in the Indian Peninsula, and relict Peninsular structural grains in the Himalayan rocks[1,28] point to an inheritance of the Indian shield elements by the Himalaya. The suggestion that the Himalaya is essentially a Hercynide fold belt[29] is also relevant in this connection. It is likely that Hercynian distension caused either opening or widening of the embryonic Tethys and that during this process rifting of the northern edge of India took place, so that a cluster of microcontinents came into existence between the main continents of India and Tibet (Fig. 2). This distension was responsible for the Gondwana graben system in India and Upper Palaeozoic sedimentary-volcanic sequence in Tibet. As this process continued, the oceanic separation between the microcontinent and the Tibetan shield increased to give rise to what may be designated as the Mesozoic Tethys. In the initial stages, the Himalayan basin[30] was a rift system and an embryonic ocean, which as indicated by Upper Palaeozoic basic
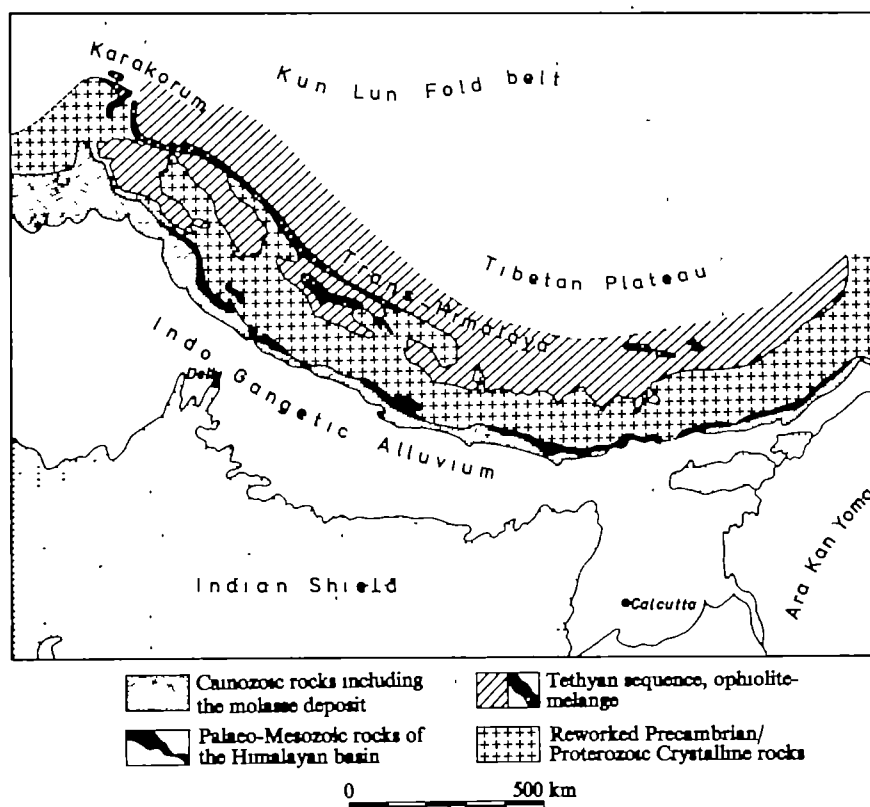
Fig. 1 Simplified geological map of the Himalaya (after ref. 1). Note extent of crystalline rocks in the core.

Cainozoic rocks including the molasse deposit

Palaeo-Mesozoic rocks of the Himalayan basin

Tethyan sequence, ophiolite-melange

Reworked Precambrian/Proterozoic Crystalline rocks

0   500 km
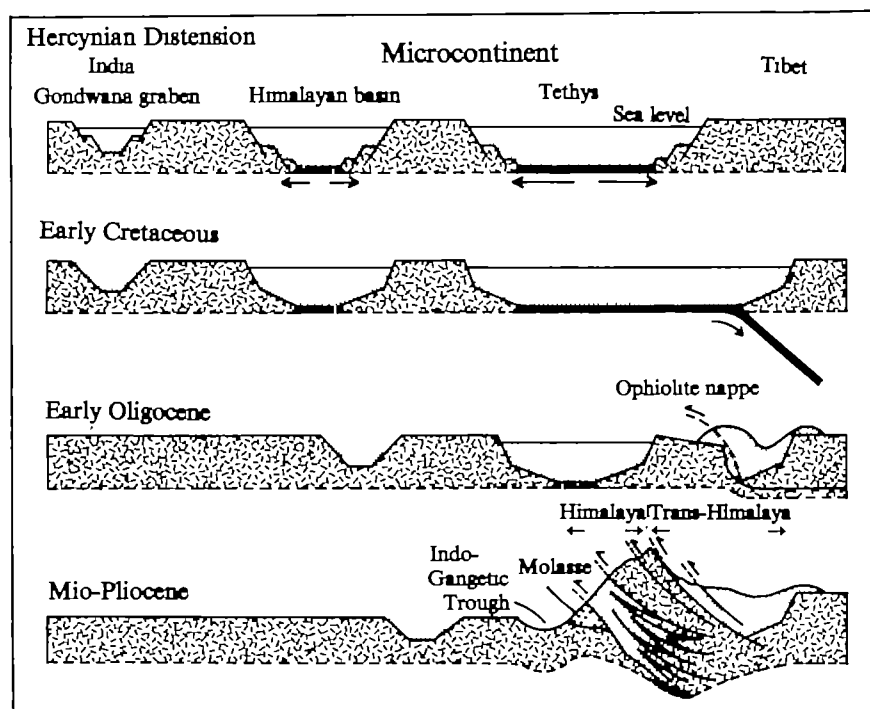
Fig. 2 Plate tectonic cartoon showing the
evolution of the Himalaya with the
involvement of microcontinent.

volcanics in Kashmir and Eastern Himalaya, contained mantle
reaching fractures[12]. Major and trace element geochemistry
of these volcanics from a part of the Eastern Himalaya (un-
published data of the author) suggests that they are transitional
between ocean floor and 'within plate'[21] basic rock suite. Most
probably, continued rifting and distension caused appreciable
thinning of the continental crust and at a late stage, formation
of an intermediate or oceanic crust in this basin. Rift tectonics
are amply demonstrated by lateral inconsistency and temporal
heterogeneity in Upper Palaeozoic and Mesozoic stratigraphy
which ranges from marine, platform to even continental types.
Local intermingling of Gondwanic fauna with the Eurasian
types in the sediments of the Himalayan basin would suggest
that the microcontinent did not behave as a solid barrier
between the Tethys proper and the Himalayan basin. There
might have been a few isthmuses.

In late Mesozoic time the Tethys began to shrink in con-
sequence of northward drift of the Indian plate which carried
along with it the microcontinent. In early Cainozoic period the
northern edge of the microcontinent collided with the Tibetan
shield along the Indus–Tsangpo suture. The Tethyan sediments
were squeezed up and ophiolite-melange exotic blocks[1] were
thrust to the south which now occur in association with the shal-
low marine sediments deposited at the northern edge of the
microcontinent. Either a part of this microcontinent or similar
continental blocks in the Tethys might have been thrust below
Tibet to cause thickening of the continental crust in the latter.
Because the orogeny had a south-directed polarity[11], later in the
collisional sequence the northern edge of the Indian shield col-
lided with the microcontinent, probably after the suturing of the
Tibetan shield and the microcontinent. As a result, the micro-
continental basement rocks, reworked during this orogeny,
were thrust over the folded and dislocated cover sediments of
the Himalayan basin to the south. As indicated by the positive
gravity anomaly, the oceanic slab of this basin probably lies
slung beneath the mountain chain. It also seems likely that
underthrusting of this slab may be linked with the generation
of Cainozoic granite-granodiorite intrusive rocks in the
Axial zone of the Himalaya, which are difficult to explain in
plate tectonic models where direct collision between India
and Tibet is postulated. The orogenic polarity is exhibited by
the Plio–Pleistocene Siwalik molasse deposits which are con-
tained in the linear trough at the down-buckled northern edge
of the Indian shield, and which are deformed and overridden

by the allochthonous Palaeo–Mesozoic rocks, and in some
cases, even by the Precambrian–Proterozoic rocks. The bi-
phasal collision scheme suggested here is supported by palaeo-
magnetic evidence[23] provided by DSDP cores which indicates
a faster rate of northward movement of India from the Cre-
taceous to the Middle Eocene (the phase of the disappearance
of Tethys and the collision of Tibet with the microcontinent),
and a subsequent slower rate from early Oligocene (the phase
of interaction between the Indian shield and the microcontinent).
The Indus–Tsangpo suture and the Trans–Axial thrust have
been virtually inactive in recent times, but the northerly thick-
ened sediments in the Indo–Gangetic trough[24], the neotectonic
features and the seismic pattern[10,24] in the Himalaya suggest
that the main movements are localised in the microcontinental
domain, along the Main Central Thrust, and in the fold–thrust
belt to the south.

It therefore seems appropriate to conclude that the Himalaya
might indeed represent reconstituted and 'digested' upthrust
microcontinental blocks which tectonically rest over the
deformed cover rocks, deposited in the Himalayan basin.
The Tethys, if it ever existed as a very wide ocean, as normally
postulated, had little direct role in controlling the stratigraphic
composition of this tectogene.

I thank Dr B. A. Sturt for comments on the manuscript.

SUBIMAL SINHA ROY*

*Geological Survey of India,
27 Chowringhee Road,
Calcutta 700013, India*

*Present address: Geologisk Institutt, Universitetet i Bergen, Avd. A., J. Frielesgt 1,
5014 Bergen, Norway.

1 Gansser, A, *Geology of the Himalaya* (Interscience, John Wiley, New York,
     1964).
2 Saxena, M. N., *Tectonophysics*, 12, 433–448 (1971).
3 Pande, L. C , and Kumar, S., *Geol. Rdsch.*, 63, 539–548 (1974).
4 Powell, C. McA., and Conaghan, P. J., *J. Geol.*, 83, 403–406 (1975).
5 Saxena, M. N., Pande, L. C., and Surendar Kumar., *J. Geol.*, 83, 399–405 (1975).
6 Wadia, D. N., *Publ. Cent adv. Study Geol* , 3, 4–8 (1966).
7 Mehdi, S. H., Kumar, G., and Prakash, G., *Himalayan Geol* , 2, 481–501 (1972).
8 Sinha Roy, S., *Abstr. Seminar on Tectonics and Metallogeny of S.E. Asia and
     Far East, World Map Commission* (1974).
9 Gansser, A., *22nd Int geol Cong.*, 11, 387–399 (1964).
10 Vakliya, K. S., *22nd Int. geol. Cong*, 11, 283–311 (1964).
11 Petrushevsky, B. A., *Tectonophysics*, 11, 29–60 (1971).
12 Sinha Roy, S , in *Geodynamics of the Himalayan region* (edit by Gupta, H. K.),
     156–166 (National geophysical Research Institute, Hyderabad, 1973).
13 Gansser, A., *Eclog geol. Helv* , 59, 831–848 (1966).

14 Dewey, J. F., and Bird, J. M., *J. geophys. Res.*, **75**, 2625–2647 (1970).
15 Dewey, J F, and Burke, K. C. A, *J. Geol.*, **81**, 683–692 (1973).
16 Le Pichon, X., Francheteau, J., and Bonnin, J., *Plate tectonics* (Elsevier, Amsterdam, 1973).
17 Powell, C McA., and Conaghan, P. J., *Earth planet. Sci. Lett.*, **20**, 1–12 (1973).
18 Jain, S. P., and Kunwar., *Nature*, **227**, 829 (1970).
19 Dewey, J. F., Pitman, W. C, Ryan, W. B. F., and Bonnin, J., *Bull. geol. Soc. Am.*, 84, 3137–3180 (1973).
20 Fitch, T. J., *J. geophys. Res.*, 75, 2699–2709 (1970).
21 Desio, A., and Marussi, A., *21st Int. geol. Cong.*, 2, 156–167 (1960).
22 Bordet, P., in *Geodynamics of the Himalayan region* (edit. by Gupta, H. K), 148–155 (National geophysical Research Institute, Hyderabad, 1973).
23 Crawford, A. R., *Geol. Mag.*, 111, 369–383 (1974).
24 Veevers, J J, Powell, C. McA., and Johnson, B D., *Earth planet. Sci. Lett.*, 27, 383–387 (1975).
25 Meyerhoff, A A., *J. geophys. Res.*, 76, 1–51 (1970).
26 Sinha Roy, S , *Himalayan Geol.* (in the press).
27 Sinha Roy, S., *J. geol Soc. India*, 14, 31–39 (1973).
28 Auden, J. B , *Rec. geol. Surv. India*, 69, 123–167 (1935).
29 Ray, D K , *Geotectonics.*, 2, 3–14 (1966).
30 Pacha, G., *23rd Int. geol. Cong.*, 3, 161–174 (1968).
31 Pearce, J. A., and Cann, J. R., *Earth planet. Sci. Lett.*, 19, 290–300 (1973).
32 Gansser, A , in *Geodynamics of the Himalayan region* (edit. by Gupta, H. K.), 97–103 (National Geophysical Research Institute, Hyderabad, 1973).
33 Blow, R. A., and Hamilton, N., *Nature*, 257, 570–572 (1975).
34 Shastri, V. V., Bhandari, L. L., Raju, A T R., and Dutta, A. K., *J. geol. Soc. India*, 12, 222–233 (1971).
35 Chowdhury, H. K., in *Geodynamics of the Himalayan region* (edit. by Gupta, H. K.), 59–71 (National Geophysical Research Institute, Hyderabad, 1973).

The sea level in the last interglacial

^{230}Th/^{234}U dating of coral terraces at 2–9 m above sea level on 'stable' oceanic islands has suggested that the sea level during the last interglacial, 120,000–140,000 yr BP (refs 1–3) was slightly higher than at present. However, accurate dating of interglacial shoreline facies on continental coasts far removed from plate boundaries has been largely neglected (the west coast of Australia[1] and Florida[4] are exceptions). Attempts have been made to date interglacial shoreline deposits from SE Australia using molluscs[5], but this has proved unreliable[6]. The recent discovery of buried corals at two localities on the E coast of Australia has provided an opportunity to determine more precise ages by ^{230}Th/^{234}U dating.

Multiple bay barriers composed of siliceous sand are a common geomorphic feature along the central and northern coasts of New South Wales. Two distinct barrier systems, the Inner Barrier and the Outer Barrier, have been recognised[7]. The Outer Barrier system is a Holocene feature, that has developed since sea level reached its present position ~ 6,000 yr ago[8]. The Inner Barrier system extends intermittently for almost 1,000 km along the coast north of Newcastle. Marine facies of this system occur at elevations of 2–8 m above present mean sea level, as determined by seams of heavy mineral and primary sedimentary structures formed on high energy, open coast beaches (tidal range ~ 2 m at spring tides). Lagoon facies are locally well developed and achieve maximum elevations of 4–6 m above mean sea level. There is no evidence of coastwise tilting of either facies in New South Wales.

The corals dated for this study come from two localities in association with the Inner Barrier, one near Newcastle (latitude 33°45′S; longitude 151°133′E), and the other near Evans Head (28°04′S; 153°27′E). In both cases the corals are in close proximity to bedrock promontories. Two coral species

from Newcastle were collected from an excavation 7 m below MSL at the base of a marine transgressive sand unit within the Inner Barrier. Corals from Evans Head were collected from an exposure behind the Inner Barrier on the south bank of the Evans River within present tidal limits. One specimen, identified as *Pocillopora damicornis*, was in its growth position, while others (*Acropora* sp., *Favites* sp.) seem to have been only slightly reworked. The corals and associated molluscan fauna suggest deposition in a shallow quiet-water lagoon, open to the sea.

The radiochemical data and ages of the corals are shown in Table 1. The corals from beneath the Inner Barrier near Newcastle give consistent ages of 143,000 and 142,000 yr BP (±12,000). The corals from behind the Inner Barrier of Evans Head gives ages ranging from 112,000–127,000 yr BP; the *in situ* coral has an age of 118,000±9,000 yr BP These dates confirm that the Inner Barrier was formed during the last interglacial. Of the five samples from Evans Head two are considered to be unreliable even though their ages agree reasonably well with the others. One coral (74630127) has a high uranium concentration and a high ^{230}Th/^{234}U activity ratio. The other coral (74630128) has a ^{234}U/^{238}U activity ratio greater than that of modern corals, and it also has a very high ^{230}Th/^{232}Th activity ratio.

The corals from Evans Head are ~ 20,000 yr younger than those from Newcastle. Evidence from New Guinea[9,10], Jamaica[11], and Timor (J. Chappell, personal communication), suggests that there were two stands of high sea level during the last interglacial at ~ 120,000 and 135,000 yr BP, with possibly a minor regression in between. The stratigraphy and morphology at the two coral sites in New South Wales indicate that the coral deposits near Newcastle could have been formed during a marine transgression before the 135,000-yr BP high sea level, while the Evans Head corals did not begin to grow until the second high sea level between 125,000 and 118,000 yr BP Although the pattern of dates from Newcastle and Evans Head are generally consistent with those established from coral terraces of tectonically uplifted areas, the precise relations between dated corals, highest sea levels reached during the last interglacial, and deposition of the Inner Barrier, are not well enough understood to establish an interglacial sea level curve for eastern Australia. The relationships between barrier and lagoon facies are, however, clear enough to point to a sea level during the final phase of the last interglacial ~ 4–6 m above present sea level.

It is difficult to determine the precise elevation of interglacial sea levels on oceanic islands. The last interglacial terrace on oceanic islands varies between 2 and 9 m above present sea level[1]. The unstable nature of many oceanic islands, and the uncertain position of dated corals with respect to sea level at the time of growth probably accounts for such differences. There is also the problem of the relationship between relative sea levels inferred from oceanic islands and those from continental coasts resulting from differences in their respective hydro-

Table 1   Radiochemistry and ages of corals from the Inner Barrier of New South Wales

Sample number	Coral species	Total U (p.p.m.)	$\dfrac{^{234}U}{^{238}U}$	$\dfrac{^{230}Th}{^{234}U}$	$\dfrac{^{232}Th}{^{230}Th}$	Age (10^3 yr)
			Newcastle Bight			
74630115A	*Goniopora lobata*	2.73±0.05	1.11±0.02	0.75±0.03	0.01	143±12
74630115B	*Blastomussa wellsi*	3.30±0.04	1.10±0.01	0.74±0.03	0.02	142±12
			Evans Head			
74630127	*Montipora* sp.	4.37±0.07	1.13±0.01	0.65±0.03	0.05	112±9
74630128	*Platygyra lamellina*	3.12±0.08	1.18±0.03	0.69±0.07	0.28	122±23
74630129	*Acropora* sp.	3.75±0.05	1.11±0.01	0.66±0.03	0.02	114±9
74630130	*Acropora* sp.	3.62±0.07	1.13±0.01	0.70±0.05	0.02	127±18
74630137	*Pocillopora damicornis**	3.44±0.05	1.11±0.01	0.67±0.03	0.02	118±9

All samples are 100% aragonite.
**In situ.*

isostatic behaviour[1,11]. Nevertheless, for oceanic islands and continental coasts not located near plate margins the degree of difference may be < ±2 m. The maximum relative mean sea level for the unwarped New South Wales interglacial deposits based on elevations of lagoonal–tidal flat surfaces is 5±1 m. This compares quite well with the recent estimate of 7.6 m for the 120,000-yr-old Waimanalo shoreline on Oahu, Hawaii[6].

In New South Wales, there is no clear evidence for marine deposition older and higher than those of Inner Barrier age. The outer continental margin is subsiding at the rate of 0.01 m per 103 yr determined from seismic and bore data from eastern Australia, a figure comparable with similar continental margins elsewhere[14]. Therefore, it is likely that sea level returned approximately to the same position, or slightly lower, during each succeeding interglacial over at least the past 700,000 yr (ref. 15). In the documented spectrum of Quaternary sea level oscillations, the last interglacial appears to. have been an exceptional event in so far as it was characterised by a great ocean water volume and warm sea temperatures[15,16]. Inner Barrier elevation and faunal composition further support this conclusion that global ice volumes were smaller in the last interglacial than at present.

The authors are indebted to Drs J. Veron and J. W. Pickett for identifying corals and molluscs, and to Drs R. McLean and P. Roy for assistance in field sampling.

J. F. MARSHALL

Bureau of Mineral Resources,
Canberra, Australia

B. G. THOM

Department of Geography,
University of Sydney,
Sydney, NSW 2006, Australia

1 Veeh, H. H., *J. geophys. Res* , 71, 3379 (1966)
2 Thomson, G. M., and Walton, A., *Nature*, 240, 145 (1972)
3 Ku, T-L., Kimmel, M. A., Easton, W. H., and O'Neil, T J., *Science*, 183, 959 (1974).
4 Broecker, W. S , and Thurber, D. L., *Science*, 149, 58 (1965).
5 Gill, E. D., and Amin, B. S., *Search*, 6, 394 (1975)
6 Kaufman, A., Broecker, W. S., Ku, T-L., and Thurber, D. L., *Geochim. cosmochim. Acta*, 35, 1155 (1971).
7 Thom, B. G., *J. Proc. R. Soc. N.S.W.*, 98, 23 (1965).
8 Thom, B. G., and Chappell, J , *Search*, 6, 90 (1975).
9 Chappell, J., *Geol. Soc. Am. Bull.*, 85, 553 (1974)
10 Bloom, A L., Broecker, W. S., Chappell, J. M. A., Matthews, R. K., and Mesolella, K. J , *Quat. Res.*, 4, 185 (1974).
11 Moore, W. S , and Somayajulu, B. L. K., *J. geophys. Res.*, 79, 5065 (1974).
12 Walcott, R. J., *Quat Res* , 2, 1 (1972)
13 Chappell, J., *Quat. Res.*, 4, 429 (1974).
14 Lowe, K. B., *Earth Sci. Rev.*, 10, 37 (1974).
15 Shackleton, N. J , and Opdyke, N D., *Quat Res.*, 3, 39 (1973).
16 Veeh, H. H., thesis, University of California, San Diego (1965).

Stab initiation of explosions

MECHANICAL initiation of fast reactions in solids can be achieved by various methods: impact[1,2], friction[3], and shock[4]. Although it is now generally accepted that the initiation starts with the degradation of mechanical energy to heat, the precise mechanisms by which the energy conversion takes place are not understood clearly. I am concerned here with the initiation mechanism in a column of a compacted solid primary explosive when stabbed with a hard small angle metallic conical striker, which I show to be frictional. A model is presented, which, for the first time, gives a quantitative estimate of the maximum temperature rise generated during the impact.

The model is based on experimental investigations of the deformation behaviour of compacts of various reactive materials (silver azide, lead azide, lead azotetrazole) when hit with steel conical indenters. These indenters had semi-apex angle 15°, were polished to a surface roughness of 5 μm and had a flat tip of diameter 15 μm. The compact of a selected material (particle size up to 25 μm) was made in an aluminium cylindrical container, 4.3 mm in diameter, 4.8 mm high and 0.25 mm wall thickness, by placing a quantity of the material in the container and compressing with a tightly fitting cylindrical punch to a load
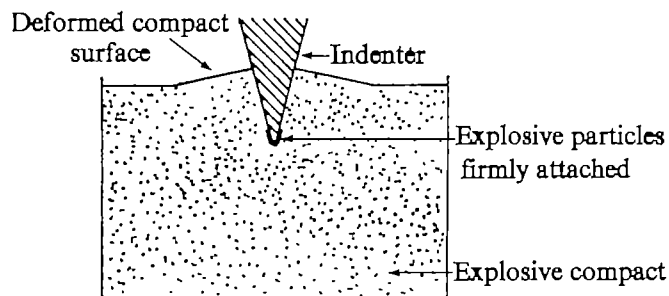
Fig. 1   Adhesion of particles to the indenter tip during penetration.

of 215 kg. Indentations were carried out using an Instron machine (crosshead speed 0.166 mm s⁻¹) and it was found that the load, L, and indenter penetration, x, were connected by

$$L = k_1 x^2 \tag{1}$$

where k_1 is a constant. An important observation was that during the indentation surface particles of the compact became firmly attached to the indenter tip and formed a thin coating on it (see Fig. 1).

Other studies on the deformation of compressed powders of inert materials with steel wedges (semi-apex angle 15°) revealed that the particle flow trajectory met the wedge surface at $\pi/2$, and this, combined with the observations described above (equation (1)), suggests that the deformation of the compact by small angle indenters can be described by plasticity theory[5] with a high value of the coefficient of friction between the indenter and the compact. This means that during the whole penetration period, the particles firmly attached to the indenter surface will rub against bulk particles with a relative velocity of $v(1 - \sin^2 15°)$, where v is the indenter velocity. On the other hand, the relative movement between the particles in other regions (that is, away from the indenter surface) of the compact will not occur for the whole period of the indentation. Moreover, as the thermal conductivity of the particles is much smaller than that of the indenter, the temperature generated at the metal/particle interface will be much smaller than at the particle/particle interface. Therefore, during indentation the maximum temperature rise will occur at the surface of the particles attached to the indenter.

Fig. 2   Calculated temperature rise against indenter penetration graphs when two conical indenters of different masses but having the same initial kinetic energy impact a column of compacted lead azide. *a*, Indenter mass 0.145 kg; velocity 0.5 m s⁻¹; *b*, Indenter mass 0.290 kg; velocity 0.353 m s⁻¹. Both indenters come to rest at a penetration of 1.16 mm. Note that the steady-state temperatures are reached well before the indenters come to rest.
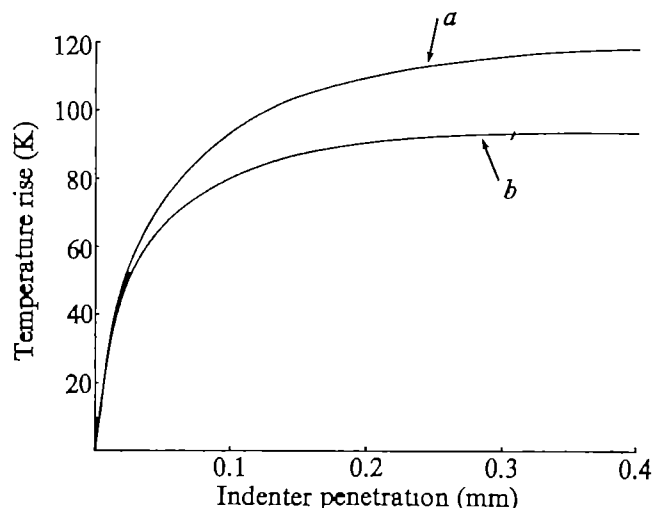
To calculate the magnitude of the maximum temperature rise it is assumed that the particles attached to the indenter tip rub against a 'continuous plane surface' provided by the particles in the bulk. Now consider the case of one of the particles attached to the indenter. The frictional heat produced per second per unit area is

$$\frac{\mu v (1 - \sin^2 15^\circ) w_p}{A}$$

where μ is the coefficient of friction, A the area of contact, and w_p the normal force on the particle. For a compact, the normal force w_p for a particle of a given size can be calculated by assuming that the normal force on the surface of the indentation is uniform. The temperature rise can then be calculated by using the method of Blok[6] (any temperature rise from chemical reactions is neglected). If the contact area is circular of radius R, and q_1 and q_2 are the fractions of heat going into the particle and the plane surface, respectively, (that is, $q_1 + q_2 = 1$), the temperature rise ΔT at the centre of the contact area after a time t is given by

$$\Delta T = \frac{4}{\pi^{\frac{1}{2}} \rho c} \int_0^t \frac{q_1 \mu v (1 - \sin^2 15^\circ) w_p (1 - \exp(-R^2 \rho c / 4t))}{2A(4\lambda t / \rho c)^{\frac{1}{2}}} dt \quad (2)$$

where ρ is the density, λ the thermal conductivity, c the specific heat and

$$q_1 = \left[1 + \sqrt{\left(\frac{\pi}{2} \frac{v(1 - \sin^2 15^\circ) R \rho c}{4\lambda} \right)} \right]^{-1}$$

Now we can calculate the maximum temperature rise when a conical striker (semi-apex angle 15°) of mass M is dropped on to a compact for which the value of the constant k_1 (equation (1)) has already been determined by experiment. The motion of the indenter will be given by

$$M \frac{d^2 x}{dt^2} = -k_1 x^2 \quad (3)$$

This can be integrated to give the indenter velocity as a function of time. Then, by using equation (2), we can calculate the temperature rise at any time (or penetration) during the impact.

As an example, the temperature rise is calculated when a conical striker, mass 0.145 kg, strikes a compact of lead azide (compaction pressure 14.8 kg mm^{-2}) with an initial velocity of 0.5 m s^{-1}. Lead azide has been chosen since all the parameters in equation (2) are known and are:

$w_p = 2.27 \times 10^{-2}$ kg (for a particle 25 μm across)
$k_1 = 3.53$ kg mm^{-2}
$c = 4.85 \times 10^2$ J kg^{-1} K^{-1}
$\lambda = 16 \times 10^{-2}$ J m^{-1} s^{-1} K^{-1}
$\rho = 4.7 \times 10^3$ kg m^{-3}
$\mu = 0.28$ (for the case of plastically deformed contact, see ref. 7)
$R = 6.4$ μm (the value of R was obtained separately by compressing a lead azide particle against a rigid surface by a load equal to the normal force on a particle of that size during the indentation of the compact).

The results of the calculations are shown in Fig. 2. Curve a is for a 0.145-kg indenter, whereas curve b is that for a 0.290-kg indenter, but with the same initial kinetic energy as before. It will be seen from Fig. 2 that the temperature rise first increases quite rapidly with penetration, but its increase then slows down, and reaches a 'steady-state' value well before the indenter comes to rest. Note that the value of the steady-state temperature rise for the 0.145-kg indenter is 118 K, whereas it is 93 K for the 0.290-kg indenter and also that, for a given penetration, the 0.145-kg indenter gives a higher temperature. This explains the observations from stab sensitivity tests of an explosive carried out with indenters of different masses that the efficiency of

initiation of reaction is higher for the indenter of smaller mass.

The steady-state temperature for the 0.145-kg indenter is 118 K plus the initial temperature of lead azide (293 K), giving 391 K; the ignition temperature of lead azide is ~ 620 K and therefore initiation of reaction is not to be expected, as is found experimentally. The same impact conditions are, however, just sufficient to cause initiation of reaction in lead azotetrazole whose ignition temperature is 493 K. The coefficient of friction for lead azotetrazole is twice that for lead azide, but the other physical properties of the two materials are similar. The estimated steady-state temperature is ~ 530 K. Therefore we expect a higher ignition probability for lead azotetrazole than for lead azide.

Our model also predicts that reaction initiation should occur at the indenter tip after the indenter has penetrated a certain minimum distance (for a given initial velocity of the indenter). High speed photographic studies of initiation of reaction in compacts of lead azotetrazole support these predictions, and a detailed account will be published elsewhere.

I would like to thank Professor D. Tabor, Dr J. E. Field and Dr J. T. Hagan for discussion and comment. The work was supported by the Procurement Executive, Ministry of Defence.

M. M. CHAUDHRI

PCS, Cavendish Laboratory,
Madingley Road,
Cambridge CB3 OHE, UK

1 Winter, R. E., and Field, J. E., *Proc. R. Soc.*, A343, 399–413 (1975).
2 Heavens, S. N., and Field, J. E., *Proc. R. Soc.*, A338, 77–93 (1974).
3 Bowden, F. P., and Gurton, O. A., *Proc. R. Soc.*, A198, 337–349 (1949).
4 Chaudhri, M. M., *Combust. Flame*, 19, 419–425 (1972).
5 Hill, R., *The Mathematical Theory of Plasticity*, 215 (Oxford University Press, 1950).
6 Blok, H., *Inst. Mech. Engrs*, 2, 222–235 (1937).
7 Amuzu, J. K. A., Briscoe, B. J., and Chaudhri, M. M., *J. Phys. D.*, 9, 133–143 (1976).

Novel electrochemical reactor

THE past decade has seen a revival of interest in most areas of electrochemical technology because of its relevance to some of the more important problems of modern society, such as environmental pollution, scarcity of natural resources and energy conservation. Electrochemical processes are quite unique in the sense that they do not need any chemical to carry out oxidation reduction reactions, instead the energy transfer is by an inert electrode. Consequently processes are generally 'clean' and do not create undesirable waste products.

The major limitation for the widespread use of electrochemical processes has been the unsatisfactory development of electrochemical reactors, but recently[1] there has been increased activity in this area. Most electrochemical processes are limited by mass transfer rates, and attempts have been made to develop reactors with greater mass transfer capabilities, for example the fluidised bed, the packed bed and other three-dimensional particulate reactors[2,3]. The limitations of these are well documented[4,5], especially the non-uniformity of electrode potential generated by poor electronic contact between the particles.

We have developed a novel electrochemical reactor based on packed carbon fibres as the working electrode. The primary design characteristic in this reactor is the increase of mass transfer rates while simultaneously having a controllable and uniform electrode potential over the entire electrode surface. The carbon fibres, which are produced by the thermal decomposition of polyacrylonitrile, have very favourable properties as electrode materials because of their high electrical conductivity and hard vitreous surface. They also show favourable hydrogen and

oxygen overvoltage characteristics with a working range of −1.4 to +1.2 V SCE in neutral solution. This gives a wide window which is especially important in the anodic range. The most outstanding property of carbon fibres for our use is their enormous surface area. Fibres are typically 5–10 μm in diameter so that 1 g of carbon fibre has a surface area of 2.6×10^6 cm (ref. 6). This is ∼ 1,000 times larger surface per unit volume than other types of particulate reactor.

As a result of the increased surface area, the mass transfer rates in this type of reactor approach those achieved by some heterogeneous catalytic reactors; as a consequence it should have immediate impact on a wide range of electrochemical processes which were hitherto economically and technologically unfeasible. As an example we can cite an application of environmental importance, the electrowinning of trace toxic metals from effluent.

The reduction of Cu^{2+} was taken as an experimental reaction, and it was found that for a 100 p.p.m. solution at a flow rate of 3.0 ml min⁻¹ a 99.3% conversion was achieved in a single pass through the reactor with a 12-s residence time. This compares to a residence time of 1–4 h for 95% conversion in packed bed and parallel plate reactors. This performance is several orders of magnitude greater than anything achieved by other reactor designs, the major improvements in the mass transfer rates arising from (1) the large surface area of fibres, (2) the low porosity and low adsorption characteristics of the fibres and (3) the fluid flow profiles in the reactor.

Scale up studies of the reactor are currently in progress, keeping the ratio of electrode area to reactor volume constant, and reactors of 10 l h⁻¹ capacity are now functioning. The optimum bed depth of carbon fibres is 1 cm. Although the primary design consideration was for mass transfer controlled reactions, it has been found advantageous as well to carry out kinetically controlled reactions in this reactor. Various electrochemical processes are being investigated using this reactor and in all of them the major advantages are:

(1) The reactor can operate at lower current densities without corresponding decrease in yield, resulting in a sharp decrease in the cost per unit of production.

(2) It is possible to operate reactions with high yield at realistic time scales.

(3) The decrease in capital cost because they can be operated in a single pass process by having appropriate design parameters.

This should lead to substantial cost saving for all processes. Details of further work with this reactor will be reported in later publications.

Bernard Fleet

Department of Chemistry,
Imperial College,
London SW7 2AY, UK

Sankar Das Gupta

Barringer Research Limited,
304 Carlingview Drive,
Rexdale, Ontario,
Canada M9W 5G2

1 Houghton, R. W., and Kuhn, A. T., *J. appl. Electrochem.*, **4**, 173 (1974).
2 Goodridge, F., *Chem. Process Engng*, **49**, 93 (1968).
3 Bennion, D. N., and Newman, J., *J. appl. Electrochem.*, **2**, 113 (1972).
4 Armstrong, R. D., Brown, O. R., Giles, R. D., and Harrison, J. A., *Nature*, **219**, 94 (1968).
5 Smith, D. H., *Chem. Process.*, **4**, (September 15, 1969).
6 *Proc. Int. Conf. Carbon Fibres* (The Plastics Institute, 1974).

Equilibrium bicontinuous structure

BICONTINUOUS partitioning of a volume is produced by inscribing a continuous, orientable surface of positive genus without self-intersection. This divides the volume into two multiply connected, interpenetrating subvolumes, each of them physically continuous (mathematically connected). Neither on a line nor in the plane is there an analogue. A bicontinuous structure is a bicontinuous partitioning in which each subvolume is filled with a distinct, not necessarily uniform composition or state of matter. In familiar examples one subvolume is solid or semi-solid, for example sandstone or sponge. An interspersion of two phases is bicontinuous only if each phase is connected across the specimen.

The idea put forward here is that bicontinuous structures may arise in fluids. The possibility of fluid analogues of porous media seems not to have been noticed before. The writer finds no mention of it in the literature on macroemulsions, periodic colloid structures, microemulsions, micellar solutions, mesomorphous phases and lyotropic crystals. Microemulsions, micellar solutions and mesomorphic phases are equilibrium states that typically consist of appreciable amounts of water, hydrocarbon and amphiphile[1,2]. Some display cubic symmetry, which is interpreted in terms of discontinuous submicroscopic globs of one composition dispersed in a continuum of another[3]. An alternative is a periodic bicontinuous structure. For thermally disrupted, non-periodic bicontinuous structures, leading candidates are optically isotropic, light-scattering microemulsion phases containing comparable amounts of water and hydrocarbon[1,4]; when one predominates, the other is often solubilised within discrete micellar structures of amphiphile. In microemulsion saturated with water and hydrocarbon, however, there is evidence that if discrete structures exist they are neither globular, tubular nor lamellar on the average and the continuous zone is dominated by neither water nor hydrocarbon[5]. The structural scale is beyond resolution by light microscopy and probably falls in the range 5–80 nm.

Intermolecular forces can conspire to elevate three energy effects to rule structure in this range. One is a sharp energy minimum when water and oil are everywhere segregated by a sheet of oriented amphiphilic molecules, the sheet constituting a partitioning surface, whether continuous or discontinuous. The second is a monotonic decrease of energy with decreasing area of the partitioning surface. The third is a sharp energy minimum with respect to the lattice parameter and symmetry of periodic structures, or mean scale length in partly randomised, isotropic structures. These three effects lead to a problem of constrained minimisation of area.

Fig. 1 Process for transforming a simple cubic array of spheres into a periodic bicontinuous structure of fused truncated octahedra with the same symmetry, lattice distance and volume fraction of one-half. The net area change is a decrease of 15%.
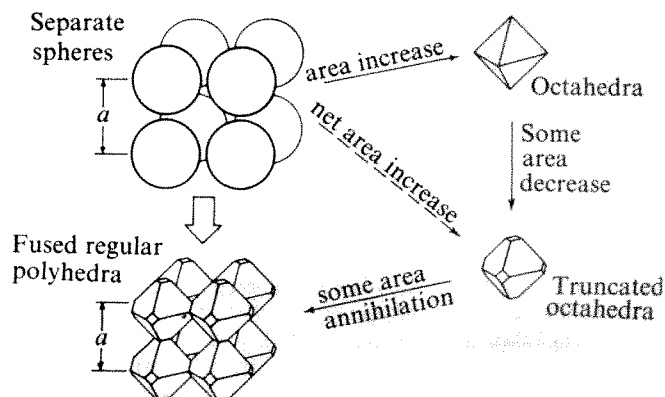
Table 1 Fused polyhedra compared with dispersed spheres in cubic arrays

	Simple cubic	Body centred	Face centred
Polyhedral subvolume	Truncated octahedra	Corner-notched	Truncated cuboctahedra
Complementary subvolume ('voids')	Truncated octahedra	Cubes and slabs	Truncated tetrahedra and truncated cubes
Volume fraction when areas are equal	0.40	0.53	0.57
Volume fraction when spheres are close-packed	0.52	0.68	0.74
Volume fraction when complementary subvolume becomes discontinuous	0.83	1	0.91
Area ratio when volume fraction is $\frac{1}{2}$	0.85	1.04	1.07
Range of volume fractions at which fused polyhedra have less total area	0.40–0.52	0.32–0.47	0.26–0.43
	0.48–0.60	0.53–0.68	0.57–0.74

Calculations show that, for cubic symmetries, there are intermediate ranges of the ratio of the two subvolumes, within which a bicontinuous structure is favoured over dispersion of either subvolume as separate spheres. Examples are fused polyhedra with the same cubic symmetry and lattice parameter as the arrays of spheres, each polyhedron having originally as many vertices as the coordination in the array. Areas at fixed volumes fractions are compared in Table 1; a process for making comparisons is pictured in Fig. 1. Notice that with body-centred and face-centred lattices there is a gap around volume fractions of $\frac{1}{2}$, in which dispersed spheres are again favoured.

Fused polyhedra represent upper bounds on partitioning area when symmetry, lattice parameter and volume ratio are fixed: the area can be further reduced by rounding edges and warping faces. The ultimate is a periodic minimal surface[6,7]. According to an unrefereed note[8], 17 intersection-free, infinite, periodic surfaces of cubic and other symmetries have been discovered. Portions of two are shown in Fig. 2 (refs 8, 9). With few exceptions[7] areas and volume fractions are unreported. Nor has their stability been studied. 'Almost periodic' and otherwise less regular minimal surfaces of high genus seem not to have been examined mathematically.

Minimal surfaces have no mean curvature and satisfy the Young–Laplace equation for menisci between immiscible bulk phases at the same hydrostatic pressure. Thus minimal surfaces of high genus might occur, if only transiently, as part of the interface in a macroemulsion, particularly in the phase-inversion range.

To account for variations in microemulsion properties Winsor[1] hypothesised that an amphiphile-dominated sheet, or partitioning surface, tends to become concave towards its hydrophilic or oleophilic side. In the concept of equilibrium bicontinuous structures this is equivalent to a fourth energy effect, a dependence of energy on curvature invariants (ref. 10, and M. L. Robbins, unpublished) or, in the extreme, a constant, non-zero mean curvature as a constraint. The minimisation problem becomes more difficult, and evidently has not been investigated even for periodic surfaces.

Equilibrium is governed by free energy, which encompasses entropic contributions besides energy effects. Thermal motions have to be accounted for in principle by identifying all energy states accessible to the system, in practice only those with energies close to that of the most probable state. Statistical thermodynamics indicates that thermal fluctuations can alter lattice distance and symmetry in periodic bicontinuous structures and that as temperature rises, periodicity and symmetry can be destroyed by the ever-greater shape excursions by the partitioning surface. On microemulsion scales these excursions are likely to cause connectivity fluctuations through pinching off and rewelding of narrow necks. As composition or temperature change and the energy minimum grows shallow compared to kT, this action can gradually yet completely destroy the continuity of one of the subvolumes, and then further reduce its mean coherence length until it is a dispersion of globs in the other subvolume. This may explain a striking feature of water–oil–amphiphile microemulsions: their continuous evolution into

a water-continuous solution of oil-swollen globular micelles on dilution with water, and into an oil-continuous solution of water-swollen inverted micelles on dilution with oil[1,2].

Bicontinuous structures may be present in mesomorphic, liquid-crystal-like phases regarded as dispersions of spheres, cylinders or lamellae. As with microemulsions, experiments are needed which discriminate unequivocally among the possibilities. Basic mathematical investigations of certain genus surfaces are indicated. Statistical physics can be brought to bear on what

Fig. 2 Examples of intersection-free, periodic, minimal surfaces, (*a*) Neovius's (ref. 9), simple cubic lattice; (*b*) Schoen's gyorid (ref. 8), body-centred cubic lattice.
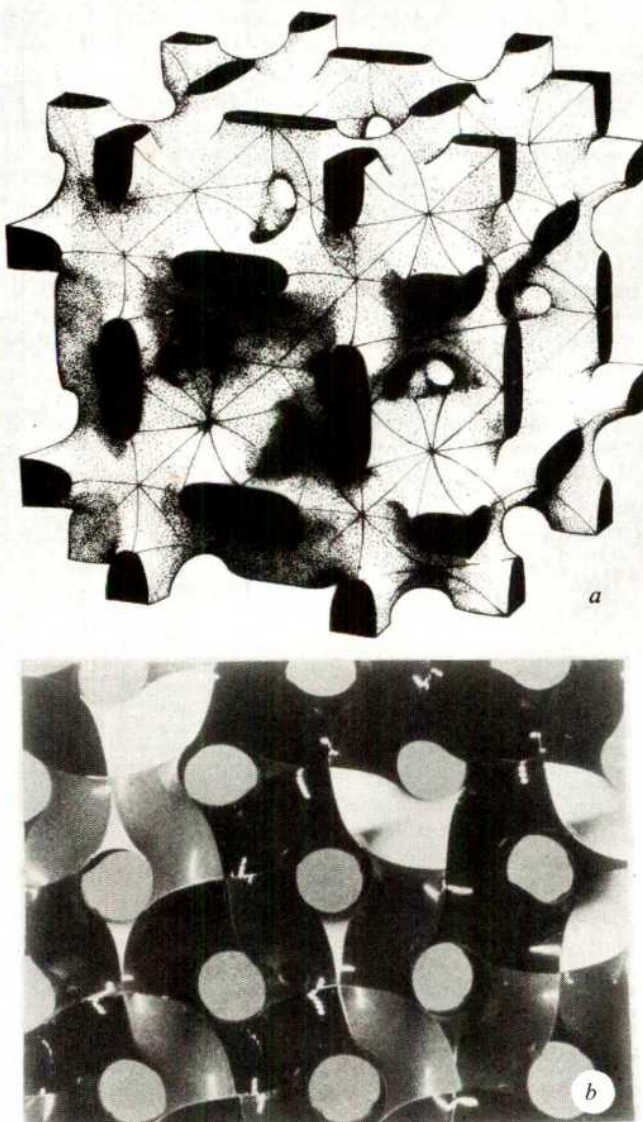
may be a state of matter related to ordinary liquids as porous media are to homogeneous solids[11].

L. E. Scriven

Department of Chemical Engineering and Materials Science,
University of Minnesota,
Minneapolis, Minnesota 55455

1 Winsor, P. A., Trans. Faraday Soc., 44, 376,382,387,390 (1948); Solvent Properties of Amphiphilic Compounds (Butterworths, London, 1954).
2 Winsor, P. A., Chem. Rev., 68, 1 (1968).
3 Ekwall, P., Mandell, L., and Fontell, L., Proc. Fifth Int. Congr. on Surface Active Substances (Barcelona), 2, 443 (1968).
4 Shinoda, K., and Friberg, S., Adv. Colloid Interface Sci., 4, 281 (1975).
5 Healy, R. N., Reed, R. L., and Stenmark, D. G., Soc. Petrol Engrs, Paper No. 5565 (presented at Dallas, Texas, 28 Sept. 1975); Soc. Petrol. Eng. J., 16 (in the press).
6 Schwarz, H. A., Gesamm. math. Abh. (Springer, Berlin, 1890).
7 Nitsche, J. C. C., Vorlesungen über Minimalflächen, Sections 276–281, 818, 882 (Springer, Berlin, 1975).
8 Schoen, A. H., NASA TN D-5541 (Washington, DC, 1970).
9 Neovius, E. R., Bestimmung zweier speziellen periodischen Minimalflächen (J. C. Frenkel, Helsingfors, 1883).
10 Murphy, C. L., thesis, University of Minnesota (1966).
11 Proc. Symp. Micellization, Solubilization and Microemulsions, Am. chem. Soc., New York, 1976 (edit. by Mittal, K. L.) (Plenum, New York, in the press).

Why be an hermaphrodite?

Many animals and most higher plants are hermaphrodites. The basic argument of this paper is that the sex habit of a species is determined by selection acting on the number of offspring produced by individuals of different types. The argument differs radically from most earlier explanations of the evolution of hermaphroditism (reviewed by Ghiselin)[1,2], although it is formally similar to a recent explanation[3] of sequential hermaphroditism, in which individuals function first as one sex and then the other.

Our fundamental assumptions are as follows. (1) Genes in a zygote can act as switches, directing development into one or other type (male, female, hermaphrodite), or, in hermaphrodites, can alter the relative allocation of resources to male and female functions. (The theory does not apply if sexual type is determined by cytoplasmic factors, as is sometimes the case for male sterility in plants.) (2) The total production of male plus female gametes by an individual is constrained to lie within a "fitness set", which cannot be altered by genetic change. It will be shown how the form of the fitness set determines the sex habit.

Consider, for concreteness, a plant species, and let m, f, and h be the numbers of male, female and hermaphrodite individuals

respectively in a population (counted at conception). For simplicity all individuals are assumed to have the same survival rate to adulthood. A male can produce N pollen grains, a female can produce n seeds, and an hermaphrodite αN pollen grains and βn seeds. The fitness set can then be represented as a graph of α against β, as in Figs 1 and 2.

Suppose the population produces R offspring. Assuming self-incompatibility, and using "fitness" to mean the expected number of offspring produced by an individual,

fitness of a male $= R/(m + \alpha h)$,
fitness of a female $= R/(f + \beta h)$,
fitness of a hermaphrodite $= R[\alpha/(m+\alpha h) + \beta/(f + \beta h)]$.

If the situation is to be evolutionarily stable, two conditions must be satisfied: (1) the fitnesses of any types actually present in the population must be equal, and (2) the values of α and β in hermaphrodites, if present, must be an "evolutionarily stable strategy," or ESS[4]; that is, the actual phenotype, $(\alpha^*\beta^*)$, of the hermaphrodites present in the population must be as fit as or fitter than any mutant phenotype $(\alpha\beta)$ lying in the fitness set.

Case 1. $h = 0$. Population dioecious. By condition (1), $R/m = R/f$, or $m = f$. This is the familiar conclusion that the primary sex ratio is 1:1 when the costs of male and female offspring are equal[5].

A dioecious population can be invaded by an hermaphrodite mutant if $\alpha/m + \beta/f > 1/m$; that is, if $\alpha + \beta > 1$. It follows that a dioecious population is stable only if the fitness set is concave (Fig. 1a). If it is convex (Fig. 1b), then only the hermaphrodite population is stable.
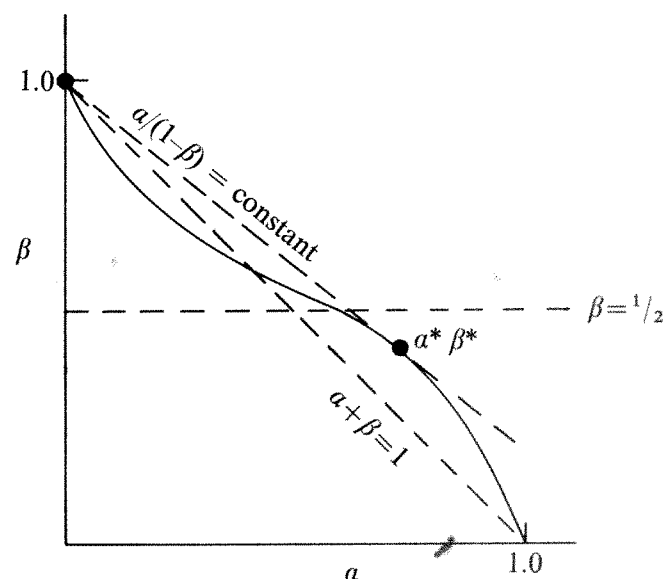
Fig. 2 An ESS which is a mixture of sexual types. If the curve is convex–concave, it may be possible for a pure sex (in this case a female) to invade. The resulting mixture is stable if (i) $\alpha^* + \beta^* > 1$, (ii) $\beta^* < 0.5$, (iii) $\alpha^*\beta^*$ is the point which maximises $\alpha/(1-\beta)$. This curve illustrates gynodioecy.

Fig. 1 a, A "fitness set" for the allocation of sex function. α is the pollen produced by an hermaphrodite, as a fraction of the total produced by a male individual. β is the corresponding parameter for a female. All possible values are assumed to be to the interior of the curve. With a concave curve, the hermaphrodite population is not an ESS, but a dioecious population is. b, Optimal resource allocation for an hermaphrodite. With a convex tradeoff curve, the hermaphrodite is an ESS. The ESS resource allocation $(\alpha^*\beta^*)$ is the point on the curve which maximises the product $\alpha\beta$.

Case 2. $m = f = 0$, $h = 1$. Population hermaphroditic. If $(\alpha^*\beta^*)$ is the phenotype of typical members of the population, and $(\alpha\beta)$ of a rare mutant, then $(\alpha^*\beta^*)$ is an ESS if the fitness of $(\alpha\beta)$ is less than or equal to that of $(\alpha^*\beta^*)$ for all points on the boundary of the fitness set. For stability against small perturbations, this requires that $(\alpha/\alpha^* + \beta/\beta^*)$ be at a local maximum when $\alpha = \alpha^*$, $\beta = \beta^*$. This is equivalent to the requirement that the product $\alpha^* \cdot \beta^*$ be a maximum (Figure 1b).

This is analogous to MacArthur's[6] conclusion that selection will maximise the product of the number of males and females in a dioecious species. If the fitness set is bounded by the line $\alpha + \beta = 1$, as will be approximately true if pollen and seed production are limited by the same resources, then the ESS is for an hermaphrodite to divide its resources equally between pollen and seeds[7].

Case 3. $h \neq 0$, $f \neq 0$, $m = 0$. Population gynodioecious. The criterion $\alpha^*\beta^*$ maximal provides a local stability criterion for a hermaphrodite population. But can a population with $h = 1$ be invaded by males or females? It can be invaded by females if $R/\beta^* > 2R$, or $\beta^* < 0.5$. Provided, however, that $\alpha + \beta > 1$ for some part of the fitness set, hermaphrodites will not be completely eliminated.

Figure 2 shows a fitness set for which the ESS is gynodioecy. Let $(\alpha^*\beta^*)$ be the stable phenotype of the hermaphrodites. Then the stable sex ratio, obtained by equating the fitness of females and hermaphrodites, is given by $f = h(1 - 2\beta^*)^8$. This implies an excess of hermaphrodites, as is in fact observed in natural populations[9]. It also implies that as the frequency of females increases, the resource allocation by the hermaphrodite shifts towards male function. This shift has also been observed[10]. The fitness of a rare mutant with a different resource allocation, $(\alpha\beta)$, is given by

$$V = R[\alpha/(m + \alpha^*h) + \beta/(f + \beta^*h)]$$
$$= R[\alpha/\alpha^* + \beta/(1 - \beta^*)]/h$$

The condition for $(\alpha^*\beta^*)$ to be an ESS is that V should be at a local maximum; which is equivalent to the requirement that $\alpha/(1 - \beta)$ should be at a maximum when $\alpha = \alpha^*$, $\beta = \beta^*$. The ESS can therefore be found by drawing a tangent as in Fig. 2.

The conditions for androdioecy, $h \neq 0$, $f = 0$, $m \neq 0$, are similar. All three types can coexist only in the artificial and unlikely case, $\alpha + \beta = 1$ exactly.

The conclusions can be summarised by saying that we would expect to find monoecy or hermaphroditism in species for which the fitness set is convex, dioecy if it is concave, and androdioecy or gynodioecy if it is concave–convex.

Are there any general reasons why fitness sets should be convex? We can offer three suggestions. (1) Low mobility: in animals, Ghiselin[1,2] has noted that hermaphroditism tends to occur in species with low adult mobility. Low mobility will tend to be associated with a convex fitness set because in such species there will be little sexual dimorphism (except perhaps dwarf males) since males will not envolve special locomotory or aggressive structures for seeking out and holding females. This means that a single individual can effectively serve both functions. Furthermore, low mobility limits male reproductive success—the number of eggs that a male can competitively fertilise (or an hermaphrodite acting as a male) should rise at less than a linear rate with resource input into male function. Note that this model assigns a different role to low mobility than the classical "low density model" as developed by Ghiselin[1,2]. Even if all the eggs produced in the population are fertilised, the convex fitness set implies hermaphroditism.

(2) Low resource overlap: in plants pollen production occurs earlier in the season, for each species, than seed maturation. The two processes therefore depend, in part, on different resources, and one would therefore expect a hermaphrodite to do better than a linear combination of male and female. In animals, brooding will imply a convex set because female expenditure will occur later than male expenditure. Ghiselin[1] has pointed out that hermaphroditism in animals tends to occur in immobile species in which individuals brood the young.

(3) Cost sharing: in insect-pollinated plants, some energy must be expended by all types on producing organs of attraction which may serve both male (pollen) and female (ovule) function. If a plant's reproductive success tends more to be limited by its ability to attract pollinators than its direct input of resources into gametes, the fitness set will tend to be convex. The conclusion depends on assumptions, which merit further investigation, on how a plant allocates resources between organs of attraction and more directly reproductive structures. It does, however, suggest that dioecy should be rarer in insect-pollinated than in wind- and water-pollinated species.

We do not claim that the selective forces associated with resource allocation are the only ones relevant to the evolution of hermaphroditism. In plants, dioecy may evolve as a mechanism to prevent self-fertilisation[11]. When associated with self-fertility, hermaphroditism may be an adaptation to situations (for example, parasitic and sessile animals; annual, monocarpic and colonising plants[12]) in which opportunities for cross-fertilisation are rare[1,2]. We have also ignored possible effects on resource allocation of factors such as sperm storage (which is common in invertebrates[1,2]) and pollen or sperm competition. Models considering these factors are now being developed and will be published elsewhere. The relative importance of these various selective forces can be determined only by comparative studies of the distribution of the sex habit, which we are undertaking.

The model outlined in this paper, and its biological interpretation, was developed independently in Utah and at Sussex.

ERIC L. CHARNOV

Department of Biology,
University of Utah

J. MAYNARD SMITH

School of Biological Sciences,
University of Sussex,
Falmer, Brighton, Sussex BN1 90G, UK

JAMES J. BULL

Department of Biology,
University of Utah,
Salt Lake City,
Utah 84112

1 Ghiselin, M. T., *Q. Rev. Biol.*, **47**, 189–208 (1969).
2 Ghiselin, M. T., *The Economy of Nature and the Evolution of Sex* (University of California Press, Berkeley, 1974).
3 Leigh, E. G., Charnov, E. L., and Warner, R. R., *Proc. natn. acad. Sci. U.S.A.*, (in the press).
4 Maynard Smith, J., and Price, G. R., *Nature*, **246**, 15–18 (1973).
5 Fisher, R. A., *The Genetical Theory of Natural Selection* (Oxford University Press, Oxford, 1929).
6 MacArthur, R. H., in *Theoretical and Mathematical Biology* (edit. by Waterman, T., and Morwitz, H.) (Blaisdell, Lexington, 1965).
7 Maynard Smith, J., in *Group Selection* (edit. by Williams, G. C.), 163–175 (Aldine, Chicago, 1971).
8 Lewis, D., *New Phytol.*, **40**, 56–63 (1951).
9 Lloyd D. G., *Genetics*, **45**, 325–339 (1975); *Heredity*, **32**, 11–34 (1974); *Heredity*, **32**, 35–44 (1974),
10 Arroyo, M. T. K., and Raven, P. H., *Evolution*, **29**, 500–511 (1975).
11 Grant, V., *Genetics of Flowering Plants* (Columbia University Press, New York, 1975).
12 Baker, H. G., *Cold Spring Harb. Symp. quant. Biol.*, **24**, 177–191 (1959).

Tip formation is regulated by an inhibitory gradient in the *Dictyostelium discoideum* slug

THE cellular slime mould *Dictyostelium discoideum* (*Dd*) has proved suitable material for investigation of pattern formation. The morphogenetic field controlling *Dd* aggregation is now uniquely well characterised[1-6]. Whether its characteristics are unique or universal is not known. Here, I help to put the slime mould into a general perspective by comparing its pattern formation with that of another well studied organism. I have used grafting experiments to investigate secondary axis regulation in the *Dd* slug. The experiments parallel the use of grafts by Wolpert and collaborators to elucidate hypostome regulation in hydra[7,8]. They show that hydra and the slug use formally similar mechanisms for axis regulation. The known features of the *Dd* aggregation control system suggest that this does not account for axis regulation in the slug. (Details of the experimental procedure are given in Fig. 1 legend.)

A tipless slug (2345) was grafted with correct polarity to a range of anterior segments. This was to test the capacity of posterior cut surfaces at various levels in the slug to inhibit tip regeneration. The result (Fig. 2a) shows a posteriorly decreasing

1 2 3 4 / 2 3 4 5
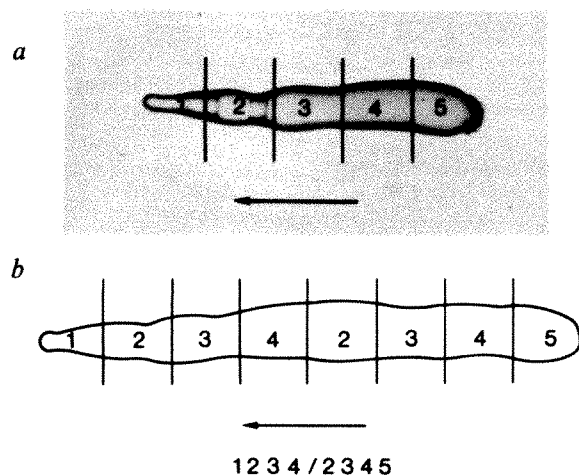
Fig. 1 Like hydra, the *Dd* slug has one obvious axis of polarity (its major axis), and a terminal organiser or dominant region (the anterior tip)[9,10]*. A second tip inserted into a slug can organise a new axis. Tip excision delays development till a new tip is regenerated. If the excised tip is replaced, regeneration is usually inhibited. Axial grafts were used to investigate the inhibition process. *a*, Large slugs (1–3 mm) were taken from growth plates and laid on non-nutrient agar[6]NC4. They were measured into five sections of equal-length (1 anterior→5 posterior) and cut at one or more section boundaries, to make desired segments (for example, 1 and 2345). *b*, Grafts were made by joining the cut surfaces of a pair of slug segments as described in the text. Anterior and posterior graft segments were always from different members of a pair of slugs approximately matched for size. Most grafts were repeated 50–100 times. Chimaeric slugs resulting from grafts usually migrated for some time before making a fruiting body.

*Hydra is bipolar, with two terminal organisers (the distal hypostome and the proximal foot). The slug is probably unipolar. Here, I compare tip regulation in the slug with hypostome regulation in hydra.

gradient in capacity to inhibit a secondary axis (linear agression: first degree coefficient = 0.22). A similar result is found for hydra[7,8,11]. The data give a reasonable fit to a linear regression line (*P* > 99.5 %) but the fit is not significantly improved for higher degree polynomial plots. Two possible trivial explanations of this result were tested and eliminated (Fig. 2*b*, Table 1).

The results above suggest that *Dd* slugs contain an axial gradient in a parameter that inhibits tip formation. If this is true, we may also expect a gradient in threshold for a response to the inhibition. The low level of inhibition at region 4/5, which is demonstrably insufficient to inhibit tip formation by region 2 (high threshold), might thus inhibit tip formation by region 5 (low threshold) in the intact slug. The results of grafts to test this prediction are shown in Fig. 2*c*. They show a posteriorly decreasing gradient in threshold, exactly as in hydra[7,8,11]. The gradient is probably nonlinear (improved fit to a second degree polynomial just significant at the 5 % level).

In hydra, secondary axis inhibition is reduced if the host organiser (hypostome) is removed just before making an axial graft[7,8]. Thus, a hydra H12/1234B56F graft makes two new hypostomes (one from each 1 region) but an H12/1234/B56F graft makes no new hypostomes and retains only the original hypostome. This result, and others, which parallel my findings in *Dd* have been interpreted (by Wolpert)[8] in terms of the following concepts.

Hypostome inhibition is proposed to be caused by a fast regulating diffusion gradient of an inhibitory substance (I). I is secreted by the hypostome and destroyed elsewhere in the hydra. The I gradient interacts with a slowly regulating gradient of similar form in a cell state variable (P). P sets the threshold for a response to I. Interactions between I and P lead to hypostome formation in a predictable way. Removal of the host hypostome causes a fall in I and enhances secondary hypostome formation at local P maxima caused by grafting. It also causes a relatively faster fall in I at the anterior end of the hydra, where

I and P concentrations are initially high, with consequent formation of an anterior hypostome. The model thus provides a rule for polarity.

Tipless slug segments were compared with their intact (tipped) homologues as inhibitors of secondary axis formation. We also tested segments lacking 12 and even 123 to seek evidence for a distributed source of tip inhibitor (Table 2). We found that removal of the slug tip never reduces secondary

Axial position of test surface
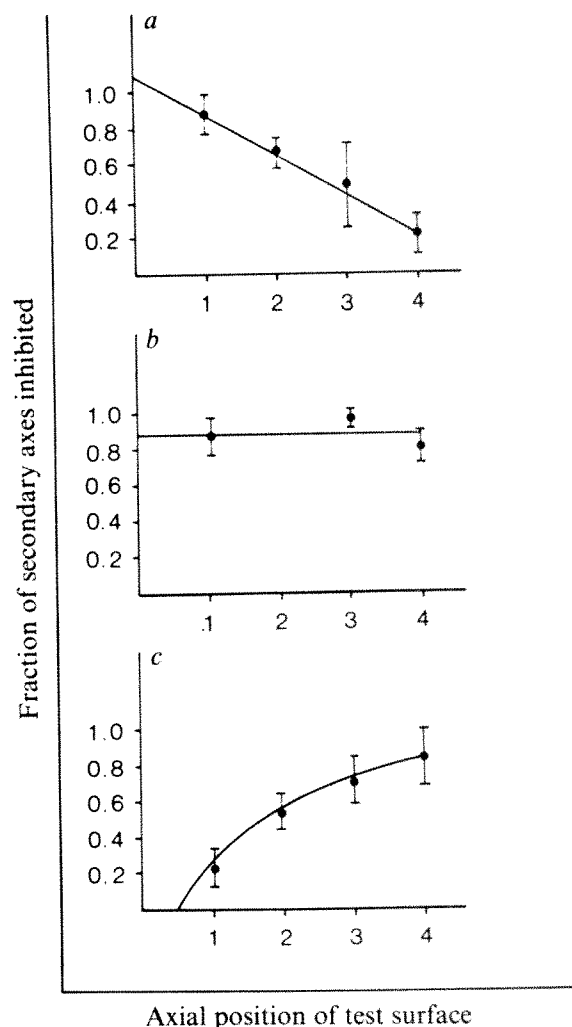
Fig. 2 Results of some grafting experiments. Ordinate, inhibition of secondary axes (axes formed by the posterior slug segment); abscissa, axial level of the cut surface in the test segment. 1, Cut between segments 1 and 2; 2, cut between segments 2 and 3 and so on. The data were obtained as results of individual experiments (8–15 grafts each). The frequency of inhibition obtained in each experiment was used as a data point to calculate regression lines, means and standard deviations. Linear regression plots are shown for 2*a* and 2*b*: a second degree polynomial for 2*c*. *a*, Test (anterior) segments 1,12,123,1234 were grafted to the standard posterior segment 2345. The plot shows a (posteriorly decreasing, approximately linear) dependence of secondary axis inhibition on the level of the cut surface in the test segment. This suggests a gradient of inhibition in the slug. Numbers of grafts for this experiment were: 1:82 2:82 3:68 4:78. *b*, A test of the possibility that the observed gradient in secondary axis inhibition reflects a gradient in adhesiveness between front cells and progressively posterior cells in the slug (with a consequent gradient in frequency of isolation of the posterior segment). The posterior segments 2345, 45, 5 were grafted to a standard anterior segment (1). These grafts show no monotonic dependence of secondary axis inhibition on level of the cut surface in the test segment (linear regression: first degree coefficient = 4.57 × 10⁻³). The possibility was eliminated. Numbers of grafts for this experiment: 1:82 3:58 4:40. *c*, Test (posterior) segments 2345, 345, 45, 5 were grafted to the anterior segment 1234. The plot shows a posteriorly decreasing gradient in the capacity of the test surface to withstand inhibition by the anterior segment. This suggests a gradient in threshold for inhibition of secondary axes in the slug. Numbers of grafts for this experiment: 1:78 2:70 3:63 4:64.

axis inhibition and neither does removal of segments back to and including segment 3. Two explanations are possible within Wolpert's type of model. (1) The *Dd* I gradient shows considerable local homoeostasis, or I diffuses very little. The tip and other distant regions are thus unimportant as a short term source of local I. P shows more homoeostasis or diffuses less than I. (2) The tip is a localised source of fast diffusing I. As in hydra, organiser (tip) removal causes a general fall in I and enhances organiser formation at local P maxima. The parameters of organiser determination and differentiation (to inhibitory competence) differ, however, between hydra and the slug. Specifically, determination and differentiation are well separated in time in hydra and not in the slug. Two slug tips developing close together may thus compete (inhibit each other). In contrast, adjacent hydra hypostomes can usually develop independently.

The second hypothesis predicts that the tip of a single chimaeric slug developing from an appropriate graft (with a tipless anterior segment) may sometimes come from the posterior

Table 1 Effect of slug length on tip inhibition

Graft	Fraction of secondary axes inhibited	No. of grafts	Mean fraction of inhibition
1234/2345	0.25	8	
	0.20	10	
	0.17	12	
	0.0	12	0.21
	0.33	12	
	0.20	15	
	0.33	9	
1234/23	0.25	8	
	0.13	8	
	0.0	10	0.15
	0.0	8	
	0.38	8	
	0.10	10	
	0.20	10	
12/2345	0.77	13	
	0.60	10	
	0.60	10	
	0.73	15	
	0.60	10	0.74
	0.80	10	
	0.79	14	
	0.88	8	
	0.88	8	
	0.75	8	

To test the possibility that the observed gradient in secondary axis inhibition arises because artificially elongated slugs (with intercalary inserts) are mechanically unstable, the grafts 1234/2345, 1234/23 and 12/2345 were compared with each other. 1234/23 gave no more inhibition of secondary axes than 1234/2345 (greater length, same anterior and posterior graft surfaces). It gave significantly less inhibition than 12/2345 (same length, anterior graft surface moved forward). The length hypothesis is ruled out and a gradient of inhibition in the slug is suggested. The figures in columns 2 and 3 are results of separate experiments.

segment. According to the first hypothesis, this result would be impossible. The prediction was tested and found true (Table 3).

The present data show that tip regulation in the *Dd* slug can be described in terms identical to those used to describe hypostome regulation in hydra. Results obtained with both organisms are well explained by assuming interactions between a fast diffusing substance (I or *r*)[8,13] and a non-diffusible or slowly diffusing substance or cell-state variable (P or *a*). In this respect, both hydra and the slug resemble a third pattern-forming system (the insect segment)[14]. These similarities suggest the existence of a general class of pattern-forming system. They indicate that insights into *Dd* pattern formation will be of general interest.

For clarity, and to emphasise similarities, I have discussed the data in terms of Wolpert's model for the regulation of hypostome formation in hydra I note that the results obtained

Table 2 Inhibition by normal and decapitated segments

Graft	Fraction of secondary axes inhibited	No. of grafts	Mean fraction of inhibition
12/2345	0.77	13	
	0.60	10	
	0.60	10	
	0.73	15	
	0.60	10	0.74
	0.80	10	
	0.79	14	
	0.88	8	
	0.88	8	
	0.75	8	
2/2345	1.00	11	
	0.88	8	
	0.63	8	
	0.88	8	0.83
	0.75	8	
	0.75	8	
	0.88	8	
1234/45	0.67	12	
	0.92	12	
	0.88	8	0.77
	0.75	8	
	0.58	12	
	0.82	11	
34/45	0.77	9	
	0.75	8	0.75
	0.63	8	
	0.83	8	
4/45	1.0	9	
	0.88	8	
	0.80	10	0.84
	0.70	10	
	0.80	8	

Tipless slug segments, and segments lacking longer anterior portions (12, 123) were compared with their intact homologues as inhibitors of secondary axis formation. None of these decapitated segments seems less effective than their homologues as inhibitors of secondary axis formation. Grafts leading to close proximity between the fronts of anterior and posterior segments gave slightly increased inhibition (2/2345 gives significantly more inhibition than 12/2345). The figures in columns 2 and 3 are results of separate experiments.

in hydra and *Dd* can be explained by at least one other type of model based on two variables[13]. There is not yet any evidence to justify commitment to a particular model.

The finding that there is an axial gradient in the capacity of a slug tip to inhibit formation of a second tip must also be evaluated in terms of what is already known about the tip as an embryonic organiser.

The tip appears first in the late *Dd* aggregate[6,9]. It is then an organiser because it is a pacemaker region, where waves of chemotactant secretion are initiated. The tip thus controls the movements of the aggregating cells. The waves have polarity (that is, propagate centrifugally), but, as far as is known,

Table 3 Fate of anterior tissue in normal and decapitated chimaeras

Graft	Fractions of grafts in which			No. of grafts
	Front tissue stays at front	Mixture of front and back tissue at front	Front tissue moves back	
12/2345	1.00	0.0	0.0	50
2/2345	0.51	1.11	0.38	47

Hypothesis 2 predicts that grafts of tipless anterior segments may give rise to chimaeric slugs in which the tip comes from the posterior segment. Grafts using tipped anterior segments should never do this. To test the prediction, we made 12/2345 and 2/2345 grafts, using a (neutral red) stained anterior segment. In 12/2345 grafts, the stained tissue always remained at the front of the chimaeric slug (there is some diffusion of dye from the stained part). In 2/2345 grafts, the whole stained piece moved backwards in 38% of cases. The slug tip was then made from white (posterior) tissue. The prediction is vindicated. The frequent posterior tip formation observed in the 2/2345 grafts implies that competition between anterior and posterior developing tips is almost unbiased in this graft. This is consistent with the strongest form of hypothesis 2 (tip as sole source of inhibitor, with little inhibition by a decapitated 2 region).

convey no scalar information. Their effectiveness as signals does not diminish with distance (*r*) from the tip[3,4,15].

The ability of an aggregate tip to inhibit formation of a neighbouring tip should work by way of entrainment of the presumptive tip tissue and its prospective field by these propagating waves. This inhibitory effect should thus be independent of *r*. The finding of a distance dependent inhibitory effect in the slug suggests that the slug tip functions differently than the aggregate tip. I suspect that this difference in function relates to the genesis of a regulative scalar pattern of differentiation in the slug[16,17]. The relevance of these findings to previous models for *Dd* pattern formation is discussed elsewhere (in preparation).

I thank Frida Vork for her help with the grafting experiments, Dorothy Parsons for typing the manuscript, and Ilse Aleven and Leen Boom for the illustrations.

A. J. DURSTON

Hubrecht Laboratory,
Uppsalalaan 8,
De Uithof,
Utrecht, The Netherlands

1 Bonner, J. T., *J. exp. Zool.*, **106**, 1–26 (1947).
2 Konijn, T., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **58**, 1152–1154 (1967).
3 Shaffer, B. M., *Adv. Morphogen.*, **2**, 109–182 (1962); *Nature* **225**, 549–552 (1975).
4 Gerisch, G., *Current Top. dev. Biol.*, **3**, 159–198 (1968).
5 Robertson, A. *et al.*, *Science*, **175**, 333–335 (1972).
6 Durston, A. J., *Devl Biol.*, **37**, 225–235 (1974).
7 Hicklin, J., Hornbruch, A., Wolpert, L., and Clarke, M., *J. Embryol. exp. Morph.*, **30**, 3, 701–725 (1973).
8 Wolpert, L., *et al.*, *Symp. Soc. exp. Biol.*, **25** (edit. by Davies, D., and Balls, M.), (Cambridge University Press, London, 1971).
9 Raper, K., *J. Elisha Mitch. Sci. Soc.*, **56**, 241–282 (1940).
10 Farnsworth, P., *J. Embryol. exp. Morph.*, **29**, 2, 253–266 (1972).
11 Webster, G., *J. Embryol. exp. Morph.*, **16**, 105–122 (1966).
12 Hicklin, J., *et al.*, *J. Embryol. exp. Morph.*, **33**, 499–510 (1975).
13 Gierer, A., and Meinhardt, H., *Kybernetik*, **12**, 30–39 (1972).
14 Lawrence, P., *et al.*, *J. Cell Sci.*, **11**, 815–853 (1971).
15 Nanjundiah, V., *J. theor. Biol.*, **56**, 275–282 (1976).
16 Takeuchi, I., *Dev Biol.*, **8**, 1–26 (1963).
17 Gregg, J., *Devl Biol.*, **12**, 377–393 (1965).

Diet-induced alterations in distribution of multiple forms of alcohol dehydrogenase in *Drosophila*

IT is now well accepted that enzyme polymorphisms are widespread among organisms. Electrophoretic polymorphisms in particular have been extensively studied and discussed[1]. One enzyme, alcohol dehydrogenase (ADH) has received special attention because the relative proportions of two polymorphic forms (allozymes), Adh[F] and Adh[S], in *Drosophila*, seem to vary with environmental temperature[2,3]. The situation with regard to *Drosophila* ADH, however, is complicated by the fact that even in flies homozygous for a given electrophoretic variant or allozyme, three forms of ADH activity are observed after electrophoresis[4–6]. They are distributed in such a way that (on agar gel electrophoresis) the one which migrates most cathodally (termed ADH[5] in the Adh[F] strain) has the highest activity, but is the least stable to heat. The least cathodally migrating form (ADH[1] in the Adh[F] strain) has the least activity and is the most heat stable, whereas the third form (ADH[3] in the Adh[F] strain) is intermediate with respect to both activity and heat stability (Fig. 1*a*). These forms are termed isozymes, in order to contrast them with the allozymes described above.

We have been interested in ADH and its genes because of their potential utility in investigating gene control in *Drosophila*[7–10]. We have formulated a model to explain the mechanism of formation of the isozymes[11], according to which, the three isozymes in any one homozygous strain have the same primary structure: they are dimers of identical subunits of molecular weight 2.5×10^4 (ref. 11). The different electrophoretic mobilities of the three isozymes result from the differential, non-covalent binding of a low molecular weight, negatively charged molecule. ADH[5] has none of this molecule
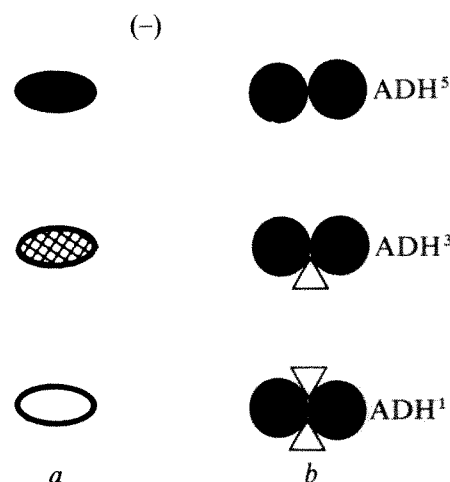
Fig. 1 Schematic representation of the isozymes of Adh[F]. *a*, Migration is towards the cathode. The most intense strain is represented by a solid oval, intermediate staining by the cross-hatched oval, and the open oval represents the least intensely stained zone. *b*, Model to explain the formation of the three isozymes. The circles represent the subunits of ADH, and the triangles a negatively charged, NAD–carbonyl compound addition complex bound non-covalently to the dimer. ADH[3] has one mol of this compound bound; ADH[1], two mol (ref. 11 and J. O'Donnell, M. S. and W. S. unpublished).

bound, but ADH[3] and ADH[1] have one and two mol bound per mol of enzyme (Fig. 1*b*). The low molecular weight molecule has been tentatively identified as an addition complex of nicotinamide adenine dinucleotide (NAD[+]) and a carbonyl compound (ref. 11 and J. O'Donnell, M. S. and W. S. unpublished) of uncertain structure. (In contrast to the epigenetic mode of formation of the isozymes, the allozymes are thought to display differences in electrophoretic migration because of changes in the primary structure of the ADH protein.)

This model makes some very specific predictions. One is that if the concentration of the addition complex is increased, the relative proportion of the isozymes should change. Specifically (in the Adh[F] strain), ADH[1] and ADH[3] should form at the expense of ADH[5]. We therefore fed Adh[F] flies a number of compounds that are known to form addition complexes with NAD[+] (ref. 12). A variety of other compounds, substrates of ADH, and some related compounds were also fed. An example of the results of these experiments is shown in Fig. 2. In this experiment, 3-hydroxy-2-butanone (acetoin) was fed to adult flies for 24 h in various concentrations. There is a decrease in
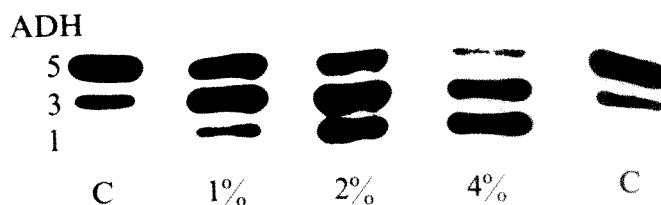
Fig. 2 Effect of acetoin in fly food on the isozymes of ADH. Acetoin was dissolved in water at concentrations indicated in the figure. (Controls, C, contained no added acetoin.) The solutions were added to four separate 2.6×8-cm vials containing dry Instant Drosophila Medium (Carolina Biologicals). Fifty adult (> 4 d old) Adh[F] flies were transferred to each of the vials and incubated at 25 °C for 24 h. The vials were sealed with a layer of Parafilm. After incubation the flies were collected, homogenised in 1 ml of 0.02 M sodium phosphate buffer, *p*H 7.5, and centrifuged at 12,000*g* for 15 min. The supernatant fluids (SFs) from these flies were then used for agar gel electrophoresis according to the procedure of Ursprung and Leone[4]. The gels were stained for 1 h at 29 °C for ADH activity using the staining procedure of Vigue and Sofer[12]. The positions of migration of ADH[5], ADH[3] and ADH[1] are shown at the left of the figure. (The photograph has been processed by a high contrast procedure in order to show the results more clearly.)

the apparent activity of ADH[5] and a concomitant increase in activity of ADH[3] and ADH[1] as a function of increasing concentrations of acetoin. (With an increased period of exposure, all the ADH activity can be shifted to the ADH[1] position.) In similar experiments, the following compounds were also found to lower the apparent activity of ADH[5] and raise that of ADH[3] and ADH[1]: 2,3-butanedione, 2,3-butanediol, 1,3-butanediol, 2-propanone (acetone), 2-butanone, 2-butanol and 2-propanol (all at 1% concentration). All these compounds possess a secondary alcohol or carbonyl group, which correlates well with the marked substrate preference of *Drosophila* ADH for secondary alcohols[4].
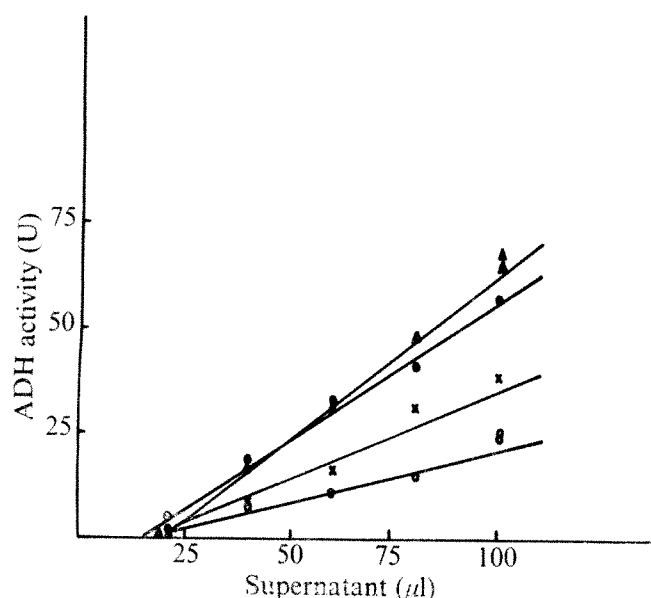
Fig. 3 Measurement of total number of ADH molecules by immunological titration of the fly supernatant fluids (SF, described in Fig. 2) shown in Fig. 2 with anti-ADH serum. Increasing amounts of fly homogenates were added to a constant volume of anti-ADH serum (prepared as described previously[10]). The abscissa indicates the volume in μl of the SF added to a constant volume (20 μl) of antiserum. After incubation for 2.0 h at 25 °C and 0.5 h at 4 °C, the mixtures were centrifuged at 12,000g for 20 min. The activity (in U per 50 μl) remaining after centrifugation is plotted as a function of μl of the original SF added. $\triangle$, SF from control (0% acetoin) flies added to 20 μl of antiserum; $\bullet$, SF from flies subjected to 1% acetoin for 24 h added to 20 μl of antiserum; $\times$, SF from flies subjected to 2% acetoin for 24 h added to 20 μl of antiserum; $\bigcirc$, SF from flies subjected to 4% acetoin for 24 h added to 20 μl of antiserum. The slopes of the four curves obtained after a linear regression analysis of the points are 0.81, 0.66, 0.50, 0.27 U per μl of the control, 1%, 2% and 4% of SFs repsectively. The x intercept, the equivalence point, is a function of the quantity of ADH-cross-reacting material. On the other hand, the slope of the line after the equivalence point represents the ADH activity per fly. Although the slopes decline threefold as a function of increasing concentrations of acetoin in the food, the equivalence points remain about the same. The equivalence points are 20, 13, 21, and 14 μl respectively.

No significant alteration in the proportions of isozymes was found after feeding with 1% butyaldehyde, *n*-butanol, pyruvic acid, glyoxal, 1,5-pentanediol, ethanol, 1-propanol or acetaldehyde. In addition, the following compounds were fed (at 1% concentration) and were found to be lethal; cyclohexanone, cyclohexanol and 2-pentanone. Many of the other compounds caused varying degrees of mortality, but enough flies remained alive for electrophoretic analysis. No simple relationship was observed between mortality and the ability to alter the ratios of the isozymes.

The changes observed after feeding acetoin and other effective compounds are termed "conversion." This term implies that either molecules originally destined to become ADH[5], form ADH[3] and ADH[1], or that ADH[5] molecules change into ADH[3] and ADH[1] molecules. In either case, the total ADH

activity would be expected to drop (since ADH[3] and ADH[1] have lower specific activities than ADH[5]) as a result of conversion, but the total number of molecules of ADH in all forms would remain the same. The data in Fig. 3 confirm our predictions and support the interpretation that conversion is taking place.

Two other predictions of the model have been confirmed. First, since strains homozygous for the three known allozymes each display three isozyme zones after electrophoresis, it is expected that each would show a similar response to being fed the effective compounds. Indeed, we found that in strains homozygous for the Adh[D] and Adh[S] variants, there was a change in the proportions of the three isozymes such that the less cathodally migrating forms seemed to increase in activity, whereas the more cathodally migrating forms lost activity after addition of acetoin to the food.

Second, since it was postulated that the isozymes originated from changes in ADH and not some other related enzyme, no bands should appear after the feeding of effective compounds in strains lacking ADH activity. In fact, no ADH activity was detectable after electrophoresis after feeding flies of the Adh[n1] strain 1% acetoin for 24 h.

Finally, we reared flies on sterile Sang's synthetic medium[14] to test the hypothesis that microbial flora were producing some compound that was causing the normal appearance of the isozymes. No differences in relative intensity of the isozymes were observed, however, when sterile cultures were compared with cultures kept in Sang's medium with added live yeast.

The experiments reported here are important for two reasons. First, since we predicted that some compounds that are capable of forming addition complexes with NAD[+] would alter the distribution of the isozymes of ADH, the data lend additional support to our model for the origin of the multiple forms of ADH. Second, these studies show that the electrophoretic migration of ADH can be changed greatly by events presumably occurring outside the *Adh* locus. It is conceivable that mutations may arise which affect the metabolism of these carbonyl compounds such that their concentration increases, resulting in changes in migration of ADH on electrophoresis. Thus, electrophoretic variants may occur in ADH with no change in the primary structure of the enzyme. For population geneticists working with this and similar enzymes, it would seem advisable, where possible, to carry out a genetic analysis of any new variants to show that the polymorphism maps close to the structural locus of the enzyme.

The three isozymes also differ significantly in stability[15], and any alteration in the ratios of activity of the three forms could cause a shift in the heat stability of total ADH extracted from flies. Since it has been reported that natural populations are heterogeneous with respect to the heat lability of some enzymes[16], it would be reasonable in the case of ADH and related enzymes to check for differences in the activity ratios of the isozymes. Here again mutations at loci outside of *Adh* which affect the amount of NAD or carbonyl compound could very well influence the stability of the enzymes without affecting its primary structure.

MARCIA SCHWARTZ*
WILLIAM SOFER

Department of Biology,
Johns Hopkins University,
Baltimore, Maryland 21218

1 Lewontin, R. C., *A. Rev. Genet.*, **7**, 1 (1973).
2 Vigue, C. L., and Johnson, F. M., *Biochem. Genet.*, **9**, 213–227 (1973).
3 Pipkin, S. B., Rhoades, C., and Williams, N., *J. Hered.*, **64**, 181–185 (1973).
4 Ursprung, H., and Leone, J., *J. exp. Zool.*, **160**, 147–154 (1965).
5 Johnson, F. M., and Denniston, C., *Nature*, **204**, 906–907 (1964).
6 Grell, E. H., Jacobson, K. B., and Murphy, J. B., *Science*, **149**, 80–82 (1965).
7 Sofer, W., and Ursprung, H., *J. biol. Chem.*, **243**, 3110–3115 (1968).
8 Sofer, W., and Hatkoff, M. A., *Genetics*, **72**, 545–549 (1972).

9 O'Donnell, J., Gerace, L., Leister, F., and Sofer, W., *Genetics*, 79, 73–83 (1975).
10 Schwartz, M., and Sofer, W., *Genetics* 83, 125–136 (1976).
11 Schwartz, M., Gerace, L., O'Donnell, J., and Sofer, W., *Third Int. Conf. Isozymes, I. Molecular Structure*, 725–751 (Academic, New York, 1975).
12 Everse, J., Zoll, E. C., Kahan, L., and Kaplan, N. O., *Bioorg. Chem.*, 1, 207–233 (1971).
13 *The Merck Index*, (Merck and Co., 1968).
14 Doane, W. W., in *Methods in Developmental Biology* (edit. by Wilt, F. H., and Wessells, N. K.), 219–244 (Crowell, New York, 1967).
15 Grell, E. H., Jacobson, K. B., and Murphy, J. B., *Ann. N.Y. Acad. Sci.*, 151, 441–455 (1968).
16 Singh, R. S., Hubby, J. L., and Lewontin, R. C., *Proc. natn. Acad. Sci. U.S.A.*, 71, 1808–1810 (1974).
17 Vigue, C., and Sofer, W., *Biochem. Genet.*, 11, 387–396 (1974).

A

a b c d   *a b c d*   *a b c d*   *a b c d*   *a b c d*
0 min   5 min   10 min   15 min   20 min

B

a b c d   *a b c d*   *a b c d*   *a b c d*   *a b c d*
0 min   5 min   10 min   15 min   20 min

Fig. 1 Heat-sensitivity profiles of *Est-6* alleles. *A*, $Est\text{-}6^{1.00}$; *a* and *b*, $Est\text{-}6^{1.00\text{-}1}$; *c* and *d*, $Est\text{-}6^{1.00\text{-}2}$; heated for the indicated times at 59.5 °C. *B*, $Est\text{-}6^{1.08}$ and $Est\text{-}6^{1.10}$: *a*, $Est\text{-}6^{1.08}$; *b*, $Est\text{-}6^{1.10\text{-}1}$; *c*, $Est\text{-}6^{1.10\text{-}2}$; *d*, $Est\text{-}6^{1.10\text{-}3}$; heated for the indicated times at 57 °C.

Heat stability variants of esterase-6 in *Drosophila melanogaster*

It has been widely recognised that electrophoretic mobility alone does not distinguish all allelic variants of enzymes that exist in natural populations[1-4]. Recent applications of various techniques, among them examination of the heat stability of allozymes, have revealed the existence of allelic variants among allozymes with identical electrophoretic mobility[5-8]. I report here the existence and frequencies of heat-stability variants of esterase-6 in populations of *Drosophila melanogaster*, as well as an electrophoretic variant whose mobility is altered by some other, probably linked, modifier locus.

Two common alleles of esterase-6 exist in natural populations as a stable polymorphism. In addition, several rare electrophoretic variants, null alleles and heat-stability alleles have been reported[9-15]. The nomenclature of Franklin[15] has been used for esterase-6 alleles. My results show that, in the populations sampled, $Est\text{-}6^{1.00}$ is a class of two alleles which can be differentiated on the basis of heat-stability, while $Est\text{-}6^{1.10}$ is a class composed of three such alleles. Finally, $Est\text{-}6^{1.08}$ has a mobility intermediate to those of $Est\text{-}6^{1.00}$ and $Est\text{-}6^{1.10}$ when homozygous but it is indistinguishable from $Est\text{-}6^{1.10}$ when heterozygous with all other lines tested (unpublished data).

Males caught in the wild were crossed individually to females heterozygous for the marked balancer chromosome TM3[16], and a series of lines homozygous for a single third chromosome were obtained through the usual series of crosses[17]. The esterase-6 locus is located at 36.8 on the third chromosome[9]. Twenty males from each line were pooled and homogenised in 0.2 ml of 0.1 M sodium phosphate buffer, pH 6.5. Aliquots of this crude homogenate were heated for 0, 5, 10, 15, and 20 min at previously determined temperatures (57.0 °C for lines carrying $Est\text{-}6^{1.10}$, 59.5 °C for lines carrying $Est\text{-}6^{1.00}$), in a constant temperature, circulating water bath with an accuracy of ±0.1 °C. The control and heated aliquots were then subjected to starch gel electrophoresis according to the methods of Richmond[18] and scored visually for staining intensity of the esterase-6 band.

Many of the extracted chromosomes carried recessive lethal genes and such lines could not be made homozygous. In these cases, if the lethal chromosome carried the $Est\text{-}6^{1.10}$ allele, it could be examined in +/TM3 flies, for TM3 carries $Est\text{-}6^{1.00}$. Lethal chromosomes carrying $Est\text{-}6^{1.00}$ could be likewise examined by outcrossing to $Est\text{-}6^{1.10}$ homozygotes, since all wild-type progeny of such crosses would be genotypically $Est\text{-}6^{1.00}/Est\text{-}6^{1.10}$. Similar crosses involving non-lethal chromosomes, which facilitated comparison of homozygotes and F_1 heterozygotes, showed the heat stability of the relevant band
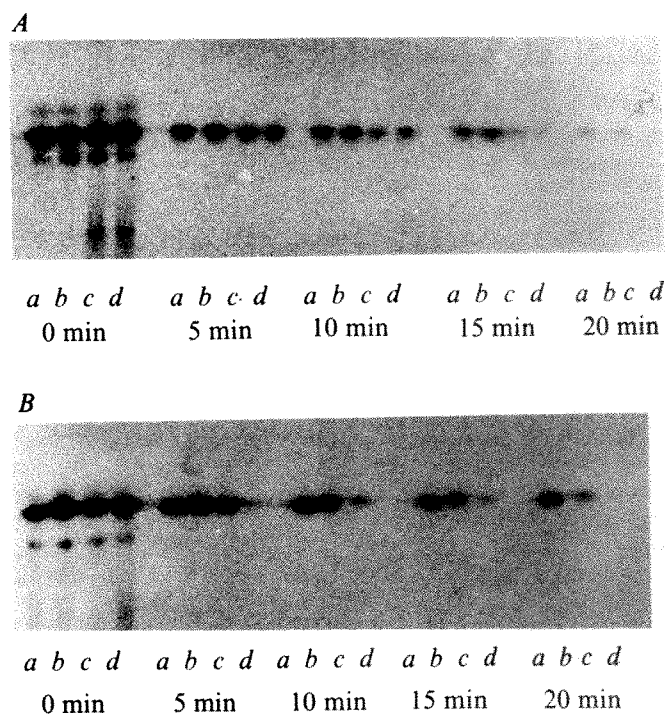
to be unaffected by being heterozygous for a second electrophoretic variant.

Figure 1*A* is a zymogram showing the two heat-stability classes found among $Est\text{-}6^{1.00}$ lines, while Fig. 1*B* shows the $Est\text{-}6^{1.10}$ heat-stability classes and the allele, $Est\text{-}6^{1.08}$, mentioned above. The class designated $Est\text{-}6^{1.10\text{-}3}$ is similar to the allele $Est\text{-}6^{F_1}$ reported by Wright and MacIntyre, while $Est\text{-}6^{1.10\text{-}1}$ is probably identical to their $Est\text{-}6^{F_2}$ (ref. 14).

It is difficult to state with certainty that the observed variation is, indeed, due to variation at the esterase-6 structural locus. The following lines of evidence are, however, consistent with such an hypothesis. (1) F_1 progeny of the crosses $Est\text{-}6^{1.00\text{-}1} \times Est\text{-}6^{1.00\text{-}2}$; $Est\text{-}6^{1.08} \times Est\text{-}6^{1.10\text{-}3}$; $Est\text{-}6^{1.10\text{-}1} \times Est\text{-}6^{1.10\text{-}2}$, and $Est\text{-}6^{1.10\text{-}1} \times Est\text{-}6^{1.10\text{-}3}$, and their reciprocals show heat stabilities intermediate to those of the parental lines. (2) F_2 progeny of the crosses $Est\text{-}6^{1.00\text{-}1} \times Est\text{-}6^{1.00\text{-}2}$ and $Est\text{-}6^{1.10\text{-}1} \times Est\text{-}6^{1.10\text{-}3}$ show segregation for heat stability not significantly different from the 3:1 ratio of stable to labile bands expected if homozygotes of the more stable allele and heterozygotes are scored as stable, while homozygotes for the less stable allele are scored as labile.

The allele designated $Est\text{-}6^{1.08}$ is exceptional. Preliminary data suggest that a second locus alters the electrophoretic mobility of the $Est\text{-}6^{1.08}$ allele, such that it is indistinguishable from $Est\text{-}6^{1.10}$. This modifier locus seems to have two allelic variants, one of which is dominant and results in this alteration. But the presence of occasional double-banded phenotypes in F_2 progeny which are segregating for the modifier locus suggests that $Est\text{-}6^{1.08}$ and $Est\text{-}6^{1.10}$ lines differ at both the structural and

Table 1 Frequencies of *Est-6* alleles observed with electrophoretic and heat-stability criteria

Location	N*	Electrophoretic alleles				Heat-stability alleles within electrophoretic classes						
		1.00	*1.08*	*1.10*	*1.15*	*1.00-1*	*1.00-2*	*1.08*	*1.10-1*	*1.10-2*	*1.10-3*	*1.15*
Bloomington, Indiana 1	38	0.757	0.108	0.135	0.0	0.135	0.622	0.108	0.0	0.027	0.108	0.0
Bloomington, Indiana 2	40	0.650	0.075	0.250	0.025	0.0	0.650	0.075	.100	0.050	0.100	0.025
McDougall, New York	22	0.682	0.0	0.318	0.0	0.091	0.591	0.0	.227	0.0	0.091	0.0
Tempe, Arizona	8	0.625	0.0	0.375	0.0	0.0	0.625	0.0	.125	0.125	0.125	0.0

*Number of chromosomes sampled.

Table 2 Latter indices of genetic distance, based on frequencies given in Table 1

Location 1	Location 2	Electrophoretic alleles			Electrophoretic and heat-stability alleles		
		δ_B	δ	$\delta_B-\delta$	δ_B	δ	$\delta_B-\delta$
Bloomington 1	Bloomintgon 2	0.396	0.378	0.018	0.505	0.478	0.027
Bloomington 1	McDougall	0.376	0.340	0.036	0.531	0.481	0.054
Bloomington 1	Tempe	0.406	0.350	0.056	0.526	0.477	0.049
Bloomington 2	McDougall	0.387	0.378	0.009	0.506	0.476	0.030
Bloomington 2	Tempe	0.402	0.387	0.015	0.484	0.472	0.012
McDougall	Tempe	0.353	0.349	0.004	0.506	0.476	0.030
Mean		0.387	0.364	0.023	0.510	0.477	0.034
Standard error		0.008	0.008	0.008	0.007	0.001	0.006

the modifier locus. Thus, for the purpose of examining gene frequencies at this locus, I have considered Est-$6^{1.08}$ to be an electrophoretic allele.

Table 1 gives the observed allele frequencies among extracted lines for the esterase-6 locus, according to geographic origin, when mobility alone, and mobility and heat stability are considered. With the exception of Est-$6^{1.16}$, which occurred only once, no allele is unique to a population, although Est-$6^{1.08}$ was found only in the Bloomington, Indiana area. This is in contrast to the results reported by Bernstein et al.[5] in which two neighbouring populations of *D. a. americana*, which shared three of four electrophoretic alleles, were found to share only one of nine stability alleles. No such radical differences in the genetic similarities of populations compared were observed in the system under consideration here.

Latter[19] has proposed a measure of genetic distance between populations that is suitable for the comparison of closely related populations whose hierarchical relationship is not known. He defines two quantities: δ, a measure of average genetic distance between individuals within populations, and δ_B, a measure of genetic distance between populations. They are defined as follows

$$\delta = 1 - \tfrac{1}{2}\sum_i (p_{i1}^2 + p_{i2}^2) - \tfrac{1}{2}\sum_i [p_{i1}^2(1-p_{i1})^2 + p_{i2}^2(1-p_{i2})^2] \quad (1)$$

$$\delta_B = 1 - \sum_i p_{i1}p_{i2} - \sum_i [p_{i1}(1-p_{i1})][p_{i2}(1-p_{i2})] \quad (2)$$

where p_{ij} is the frequency of the ith allele in the jth population. The net difference between populations is thus $\delta_B-\delta$.

Values for all pairwise comparisons of the populations sampled are given in Table 2, using both electrophoretic and heat stability classes. While application of the second criterion does increase δ and δ_B, the average net difference between populations is not increased significantly. Thus, the pattern of relative constancy of allele frequencies reported by O'Brien and MacIntyre[10] on the basis of electrophoretic mobility alone is not significantly altered by application of the heat stability criterion.

Genetic distance estimates are usually based on an averaging of several loci in populations[19]. However, Hedrick[20] has pointed out that if selective forces vary from locus to locus, such averaging will tend to give deceptive results. Therefore, the use of such an approach at the one locus studied here is valid as a measure of population differences at that locus.

Analysis of the reported data does not follow the predictions of Maynard Smith[3] and King[4], who propose that such variation is selectively neutral, and that the geographically uniform patterns of gene frequencies observed are a result of the limited number of charge states which the protein can assume. If such were the case, frequencies of hidden alleles should show a more random distribution than those of electrophoretic alleles. In the case of esterase-6 in *D. melanogaster*, no significant increase in randomness of gene frequencies is evident. This fact and these data are not sufficient, however, to indicate the action of selection at this locus.

I thank Rollin C. Richmond for guidance and counsel, and Ms Carolyn Wilhelm for technical assistance. I acknowledge

BRUCE J. COCHRANE

Department of Zoology,
Indiana University,
Bloomington, Indiana 47401

1 Lewontin, R. C., *The Genetic Basis of Evolutionary Change* (Columbia, New York, 1974).
2 Ohta, T., and Kimura, M., *Am. Nat.*, 109, 137–145 (1975).
3 Maynard Smith, J., *Nature new Biol.*, 237, 31 (1972).
4 King, J. L., *Genetics*, 76, 607–613 (1974).
5 Bernstein, S. C., Throckmorton, L. H., and Hubby, J. L., *Proc. natn. Acad. Sci. U.S.A.*, 70, 3928–3931 (1973).
6 Singh, R. S., Hubby, J. L., and Lewontin, R. C., *Proc. natn. Acad. Sci. U.S.A.*, 71, 1808–1810 (1974).
7 Singh, R. S., Hubby, J. L., and Throckmorton, L. H., *Genetics*, 80, 637–650 (1975).
8 Trippa, G., Loverre, A., and Catamo, A., *Nature*, 260, 42–44 (1976).
9 Wright, T. R. F., *Genetics*, 48, 787–801 (1963).
10 O'Brien, S. J., and MacIntyre, R. J., *Am. Nat.*, 103, 97–114 (1969).
11 Rodino, E., and Martini, A., *Drosphila Inf. Service*, 46, 139 (1971).
12 Triantaphyllidis, C. D., and Christadoulou, C., *Biochem. Genet.*, 8, 383–390 (1973).
13 Franklin, I. R., *Drosophila Inf. Service*, 47, 113 (1971).
14 Wright, T. R. F., and MacIntyre, R. J., *J. Elisha Mitchell Sci. Soc.*, 81, 17–19 (1965).
15 *Drosophila Information Service*, 47, 113 (1971).
16 Lindsley, D. L., and Grell, E. H., *Genetic Variations of Drosophila melanogaster* (Carnegie Institute, Washington, DC, 1967).
17 Wallace, B., *Topics in Population Genetics* (Norton, New York, 1968).
18 Richmond, R. C., *Genetics*, 70, 87–112, (1972).
19 Latter, R. D. H., in *Genetic Structure of Populations* (edit. by Morton, N. E.), 27–37 (University Press of Hawaii, Honolulu, 1973).
20 Hedrick, P. W., *Evolution*, 29, 362–365 (1975).

Method to test inhibitory antibodies in human sera to wild populations of *Plasmodium falciparum*

PROTECTIVE antibodies are involved in acquired immunity to malaria parasites[1]. Sera from monkeys immune to *Plasmodium knowlesi* contain antibodies which depress multiplication of the parasite *in vitro*[2] by, it is believed, inhibiting the invasion of red cells by merozoites[3,4]. Antibodies inhibitory to parasites of a specific antigenic variant are of a higher titre than inhibitory antibodies which cross-react with different antigenic variants[5]. We have attempted to extend these findings to natural human infections with *P. falciparum* in The Gambia, West Africa, where *P. falciparum* malaria is hyperendemic. The entire population is exposed to infection and an effective immunity is only acquired over 4–5 yr (ref. 6). Immunity to *P. falciparum* can be passively transferred with IgG from immune Gambian adults[7].

Gambian sera were screened *in vitro* for anti-*P. falciparum* activity as follows. Ring-stage parasites obtained from infected children by venepuncture were grown through schizogony in microtissue cultures[8] containing in replicate 150 μl of supplemented medium TC-199 and 100 μl of the following sera: (1) serum from the donor of the infected red cells; (2) Caucasian AB serum collected in the United Kingdom, and (3) serum from unrelated Gambians (matched for the ABO blood group of the infected cell donor). Parasite growth after reinvasion was measured by
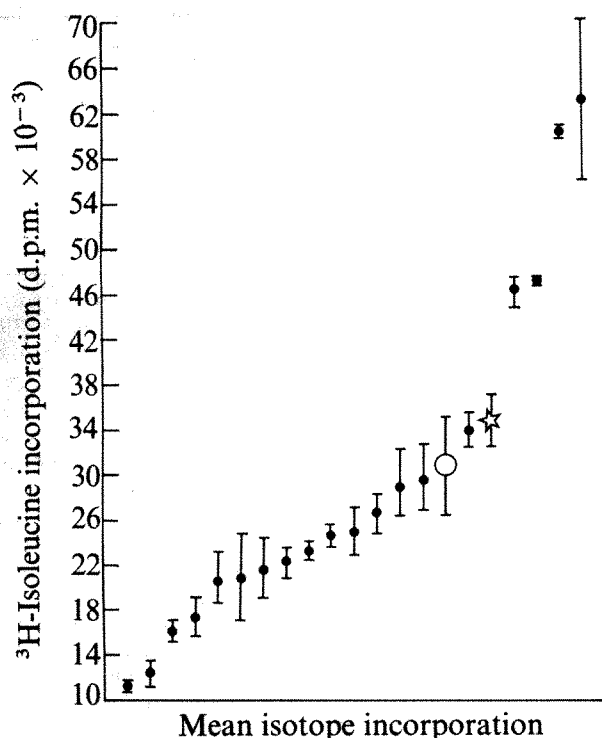
Fig. 1 Incorporation of ³H-isoleucine by *P. falciparum* parasites grown through schizogony *in vitro* in the presence of sera from 21 different donors. Isotope (specific activity 26 Ci mmol⁻¹) was added after schizogony to give 1 µCi per culture well. Cultures were collected 43 h later. Star, autologous serum; ○, Caucasian AB serum; ●, African adult serum. Bars represent range of values.

incorporation of ³H-isoleucine and parasite multiplication was assessed by microscopic examination of smears of the cultured blood stained with Giemsa stain.

In Fig. 1, the inhibitory effects on parasite growth of sera from 19 immune African adults are compared with the effects of two control sera. Most of the sera from African adults inhibited parasite multiplication and hence the incorporation of isotope in the ensuing growth period. In a few (4/19) sera, parasite multiplication and isotope incorporation were enhanced compared with controls. More than 100 sera from Gambians aged 1–65 yr have been tested with parasites isolated from many different children. In about 71% of the sera the multiplication rate of the parasites was at least 10% less than that in serum from the donor of the parasites. But multiplication was inhibited by 20% or more in only 61% of the sera tested and by 50% or more in only 31% of the sera tested. Besides this rather low degree of inhibition, variability in the amount of inhibition was a frequent feature when the same sera were tested against parasites from different children. Table 1 summarises data from an experiment in which 11 sera from Gambian adults were tested against parasites isolated from four different infected children; the individual sera inhibited parasites from some children but not from others.

These findings raise questions about antigenic diversity in *P. falciparum* parasites, as strong inhibitory activity might depend on correct matching of parasite (antigenic) type and antiserum. In addition, the specific inhibitory antibody response might be short lived, like the precipitating antibody response to certain *P. falciparum* antigens[9]. Matching of parasite population to antiserum was not controlled and there was no knowledge of the temporal relationship between them in the experiments discussed above, nor was there in the studies of Phillips *et al.*[10] and Mitchell *et al.*[11]. The remainder of this report concerns a method we have

devised to test known homologous or heterologous combinations of parasitised cells and sera.

Heparinised blood (17 U ml⁻¹), as well as serum, was obtained from infected children. The blood was cryopreserved in a final concentration of 10% glycerol in Ringer solution at −20 °C; the serum was also frozen. Each child was treated once with chloroquine (5 mg base per kg) immediately after the infected blood was collected and a second serum sample was obtained 1–3 weeks later, after catabolism of the antimalarial. When the second serum sample was collected, 43 of the 50 donor children no longer had a patent asexual parasitaemia. It was expected that in the interval between collection of the two serum samples, each child would develop a specific immune response to the parasite population of which a representative sample had been cryopreserved. Precipitating antibodies usually appear within a week or two of treatment[12]. To compare the inhibitory activities of the first (I) and second (II) serum samples, the appropriate frozen parasites were reconstituted and grown through schizogony *in vitro* in the presence of known homologous or heterologous sera.

Table 1 Distribution of 11 adult sera according to the degree of inhibition they produced when tested with four different isolates of *P. falciparum*

Percentage inhibition of parasite multiplication	Different isolates of infected red cells			
	1	2	3	4
None	3	0	11	1
Up to 20	5	1	0	4
Up to 40	3	2	0	5
Up to 60	0	8	0	1

Studies on the metabolism[13] of chloroquine in man indicate that the dose we administered should have reached the minimum therapeutic level by day 7, the earliest time that we obtained the second serum sample from any child. To check for any antimalarial activity through residual chloroquine, we carried out cytotoxic tests with the sera in microtissue cultures using *P. knowlesi* as a model system. The results indicated that there was an overall reduction in the uptake of ³H-isoleucine by *P. knowlesi* parasites grown in serum (II) samples, but this was not significant at the 5% level. The few (5/31) serum (II) samples that had a marked inhibitory effect on *P. knowlesi* did not consistently inhibit *P. falciparum*, nor were they among the samples taken

Table 2 *Plasmodium falciparum* multiplication in serum obtained before (I) and after (II) antimalarial treatment

System	Serum	cells	(I)	(II)	
Homologous	7529	7529	100	(12)	19
	7536	7536	100	(14)	11
	7540	7540	100	(14)	45
	7544	7544	100	(14)	219
Heterologous	7519	7529	102	(14)	20
	7523	7529	52	(14)	28
	7525	7529	42	(7)	26
	7519	7536	56	(14)	2
	7529	7536	113	(12)	4
	7519	7539	136	(14)	56
	7523	7539	133	(14)	122
	7525	7539	96	(7)	67
	7529	7539	157	(12)	67
	7536	7539	135	(14)	102
	7506	7540	52	(14)	17
	7531	7540	146	(14)	108

Results are expressed as percentage of multiplication in homologous serum (I). Triplicate tissue cultures each contained 10 µl of blood + 150 µl TC-199 + 100 µl of serum. Percentage multiplication was calculated from the mean results. Italic figures indicate number of days between antimalarial treatment and collection of serum (II).

Table 3 Uptake of ^{3}H-isoleucine (c.p.m. $\times 10^{-3}$) by *P. falciparum* grown in serum obtained before (I) and after (II, III) antimalarial treatment

System	Serum	Cells	AB	I	II		III	
Homologous	7525	7525	1.31	3.21	*(7)*	0.23	*(24)*	2.09
	7531	7531	ND	3.32	*(14)*	0.05		
	7544	7544	2.25	5.98	*(14)*	12.84		
	7580	7580	1.46	2.54	*(7)*	0.28		
	7584	7584	1.76	2.19	*(7)*	0.73		
Heterologous	7569	7584	1.76	4.13	*(13)*	1.08		

Duplicate tissue cultures each contained 10 µl of blood + 150 µl of medium TC−199 + 50 µl of serum + 10 µl of ^{3}H-isoleucine (1µ Ci per well). Incorporation values are given as means. Italic figures indicate number of days between antimalarial treatment and collection of serum (II) or (III). ND, Not done; AB, Caucasian AB serum.

only 7 d after antimalarial treatment. Factors other than residual antimalarial may account for this inhibitory activity. Furthermore, a few children returned 2–3 weeks after chloroquine treatment, with patent parasitaemias and in these children there presumably was no persistent chloroquine activity.

Data on multiplication of *P. falciparum* parasites in children's sera taken before (I) and after (II) antimalarial treatment are shown in Table 2. Parasite multiplication is given as a percentage of that in homologous (I) serum, so figures greater than 100% indicate enhanced parasite growth in homologous (II) or heterologous sera.

Marked inhibition of parasite multiplication occurred in three of the four examples shown of homologous serum–cell mixtures. In the fourth example (7544) parasite multiplication was enhanced in homologous (II) serum (see below). Table 2 also shows results obtained with four isolates of infected cells that were tested against heterologous sera. Again there was marked inhibition of parasite multiplication in most of the serum (II) samples (*t* test for 12 paired (non-independent) samples gave $P<0.001$). On the other hand, in heterologous (I) samples, inhibition occurred in only 4/12 cases (some cases of cross inhibition might be expected because all the children had been exposed previously to repeated malarial infection). The inhibitory activity of serum 7536 (II) in the homologous but not in the heterologous system (7539 cells) is notable. Specific inhibition is also suggested in the heterologous tests of serum 7523 (II) with cells from 7529 (inhibition) and 7539 (no inhibition).

Measurements of ^{3}H-isoleucine incorporation gave similar results. Table 3 shows that serum 7525 (III) taken 24 d after antimalarial treatment was less inhibitory than serum (II) obtained 7 d after treatment. Serum (II), however, was not inhibitory in the *P. knowlesi* cytotoxic assay. By contrast, serum 7544 (II) was inhibitory in the *P. knowlesi* assay but enhanced both parasite multiplication (Table 2) and isotope incorporation (Table 3) in homologous tests with *P. falciparum*. No explanation for this enhanced growth can be offered at present but it may possibly be a nutritional effect.

In conclusion, our results indicate that in most cases there was significant inhibition of parasite multiplication by children's convalescent sera. Both 'cross reactive' and specific inhibition reactions were observed. The possible roles of anti-parasitic antibodies, autoantibodies to red cells, and non-antibody effects in this system require further investigation, as does the blocking effect of malarial antigen released into the serum during infection or added experimentally.

Studies of their antigens[14] and isozymes[15] suggest that wild populations of *P. falciparum* parasites frequently consist of mixed types. At present there are no reliable methods either to measure the extent of this heterogeneity or to select components from it. The protocol we present here circumvents this problem and makes possible a test of the homologous inhibitory antibody response to a pure or mixed

population of parasites. The methods that we used to cryopreserve and reconstitute parasitised erythrocytes were unrefined but further work using 10% DMSO as the cryoprotectant has improved this aspect of the test. Consequently we feel that our method offers a practical approach to the analysis of anti-parasitic antibodies in human malaria infections. Cryopreservation of *P. falciparum*[16] can be used to look for other functional antimalarial immune responses, such as cell-mediated cytotoxic responses[17], directed against specific populations of the parasite at a particular time, and also to provide laboratories outside the tropical regions with infected blood, for example for *in vitro* drug screening.

We thank Dr Elizabeth Spalding and Mr A. K. Rahman for helping to select suitable malaria cases and obtain blood specimens from them. We also thank Dr R. S. Bray and his staff at the MRC Laboratories in The Gambia for laboratory facilities. R.S.P. was supported by a US Army Contract.

R. J. M. WILSON

Division of Parasitology,
National Institute for Medical Research,
London NW7 1AA, UK

R. S. PHILLIPS

Department of Zoology,
University of Glasgow,
Glasgow G12 8QQ, UK

1 Brown, I. N., *Adv. Immun.*, 11, 267 (1969).
2 Cohen, S., Butcher, G. A., and Crandall, R. B., *Nature*, 223, 368 (1969).
3 Miller, L. H., Aikawa, M., and Dvorak, J. A., *J. Immun.*, 114, 1237 (1975).
4 Butcher, G. A., and Cohen, S., *Trans. R. Soc. trop. Med. Hyg.*, 64, 470 (1970).
5 Butcher, G. A., and Cohen, S., *Immunology*, 23, 503 (1972).
6 McGregor, I. A., Gilles, H., Walters, J. H., Davies, A. H., and Pearson, F. A., *Br. med. J.*, 2, 686 (1956).
7 Cohen, S., McGregor, I. A., and Carrington, S., *Nature*, 192, 733 (1961).
8 Phillips, R. S., Rahman, A. K., and Wilson, R. J. M., *Trans. R. Soc. trop. Med. Hyg.*, 69, 432 (1975).
9 McGregor, I. A., and Wilson, R. J. M., *Trans. R. Soc. trop. Med. Hyg.*, 65, 136 (1971).
10 Phillips, R. S., Trigg, P. I., Scott-Finnigan, T. J., and Bartholomew, R. K., *Parasitology*, 65, 525 (1972).
11 Mitchell, G. H., Butcher, G. A., Voller, A., and Cohen, S., *Parasitology*, 72, 149 (1976).
12 Wilson, R. J. M., and Bartholomew, R. K., *Parasitology*, 71, 183 (1975).
13 McChesney, E. W., Banks, W. F., Jr, and McAuliff, J. P., *Antibiotics Chemother.*, 12, 583 (1962).
14 Wilson, R. J. M., McGregor, I. A., Hall, P., Williams, K., and Bartholomew, R., *Lancet*, ii, 201 (1969).
15 Carter, R., and McGregor, I. A., *Trans R. Soc. trop. Med. Hyg.*, 67, 830 (1973).
16 Pavanand, K., Permpanich, B., Chuanak, N., and Sookto, P., *J. Parasit.*, 60, 537 (1974).
17 Coleman, R. M., Renericca, N. J., Stout, J. P., Brissette, W. H., and Smith, D.M., *Immunology*, 29, 49 (1975).

Possible repair of carcinogenic damage caused by dimethylnitrosamine in rat kidney

SINCE the discovery that cells from patients with xeroderma pigmentosum, a disease associated with a high incidence of cancer, may be unable to excise ultraviolet light-induced thymine dimers from DNA[1], and the recognition of the possibility that all carcinogens react with the DNA of the organs in which tumours are induced[2], it has been thought that repair of changes produced in DNA by carcinogens

might be important in moderating the effectiveness of these compounds. Work on the carcinogenic effect of prolonged administration of various chemicals[3], however, suggests that individual carcinogenic doses are entirely cumulative in their action, and that the carcinogenic damage is permanent and not repaired[4]. If this were correct, repair of lesions produced in DNA by carcinogens would seem to have no direct relevance to carcinogenesis. Recovery from the lethal effects of radiation was established by the classic experiments of Elkind and Sutton[5] in which cultured cells were irradiated twice with different intervals between the first and second dose; if sufficient time elapsed between the first and second dose, the effects of the doses were not cumulative. We have used an analogous experimental design to study the cumulation of the effect of dimethylnitrosamine in the induction of kidney tumours in the rat. The experiment is not yet complete, but we believe we have demonstrated that a repair process exists.

A single dose of dimethylnitrosamine induces cancer in the rat kidney. When rats are fed a protein-free, high-carbohydrate diet for 7 d before receiving the dimethylnitrosamine they will tolerate a single dose which will induce cancer in every rat[6]. A few tumours are induced in other organs, mainly in the lung, but the kidney tumours predominate and the other tumours do not kill the animal before the rats are old enough for any potential kidney tumour to have become detectable.
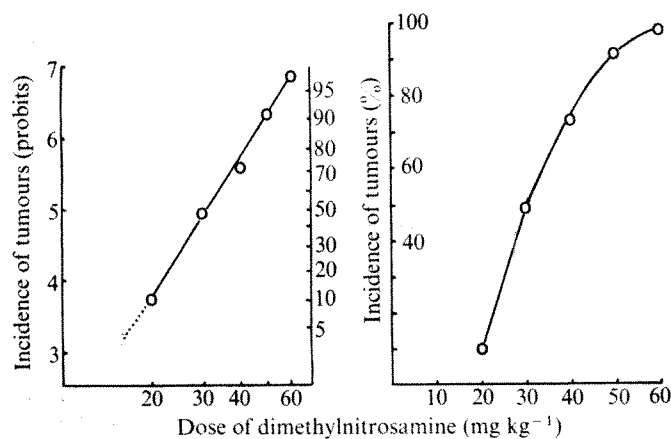
Fig. 1 Dose–response curve for the induction of kidney tumours 80 weeks after a single administration of dimethylnitrosamine. Rats were obtained from the same source and fed the same diet as in the experiment described in the legend to Fig. 2. Results are shown as *a*, incidence of tumours against dose; *b*, as the probit of incidence against the log of the dose; *b* Was used to predict the incidence which might have been expected if other doses had been given. It can be seen that the incidence predicted from a single dose of 32 mg kg^{-1} (52%) and from 16 mg kg^{-1} (5%) were very close to those obtained in the experiment shown in Fig. 2 (55% and 6%).

The dose–response curve for the induction of kidney tumours was determined in the USA between 1972 and 1974. From this curve (Fig. 1) it can be predicted that a single dose of 16 mg per kg body weight will induce kidney tumours in 5% of the rats 80 weeks after the dose was given, rising to 10% after 2 yr. In contrast 32 mg kg^{-1} will induce a 52% incidence at 80 weeks rising to 80% after 2 yr. We have estimated the rate and extent of recovery from carcinogenic damage from the incidence of tumours produced by one dose of 16 mg kg^{-1} followed after some time by another dose. If the two doses were to be given so close together that no recovery from the carcinogenic damage could occur, the two would be completely cumulative in their effect and would be expected to give the same tumour incidence as a single dose equal to the sum of the two doses, that is, 32 mg kg^{-1} giving a 52% incidence at 80 weeks, or 80% at 2 yr. If, however, repair of the damage
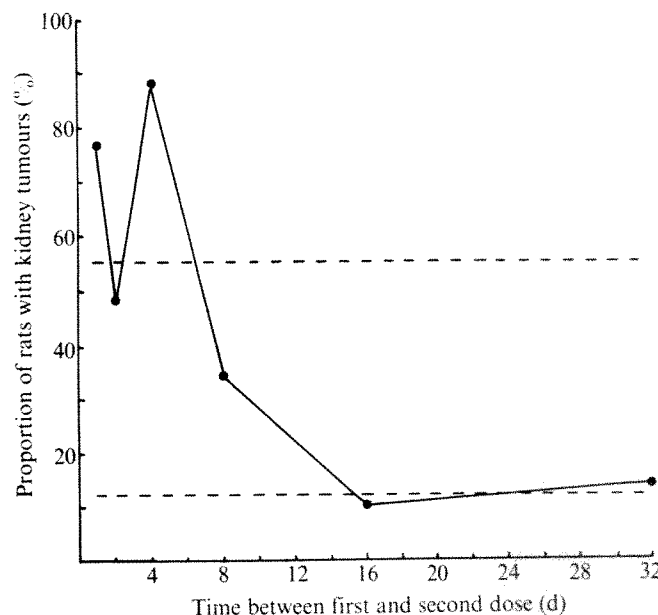
Fig. 2 Incidence of kidney tumours induced 80 weeks after a single intraperitoneal injection of dimethylnitrosamine (32 mg per kg body weight) (100 rats) a single dose of 16 mg kg^{-1} (140 rats) or by two doses, the first of 16 mg per kg body weight, the second a "nominal" (see below) 16 mg kg^{-1} 1, 2, 4, 8, 16 or 32 d later. (There were 50, 50, 60, 50, 100 and 100 rats in each group.) Male CFN rates (Carworth Farms, New City, New York) each weighing about 180 g were changed to a protein-free diet for 7 d, given an injection of dimethylnitrosamine, and then maintained on the same diet for 5 d. They were then returned to a normal rat diet (MRC 41B, Christopher Hill). The second dose was adjusted, as described in the text, so that the total amount of the kidney, when two doses were given, was believed to be the same as that resulting from a single dose of 32 mg kg^{-1}. This adjustment meant that the actual dose given (in mg kg^{-1}) 1 d after the first was 12; 2 d 13; 4 d 12.5; 8 d 26; 16 d 24 and 32 d 25. Eighty weeks after the first dose, tumours had been found in 212 of the 617 rats surviving more than 4 weeks. The upper dotted line shows the incidence produced by a single dose of 32 mg dimethylnitrosamine per kg body weight. If the doses were completely cumulative, one would expect two with the effect of 16 mg kg^{-1} to give this incidence of tumours. A single dose of 16 mg kg^{-1} gave an incidence of 6%. If the doses were not cumulative and each dose of 16 mg kg^{-1} acted independently of the other, one would expect to get twice the incidence produced by a single dose of 16 mg kg^{-1} (that is 12%) marked by the lower dotted line.

caused by the carcinogen occurs in the time between the first and second dose, a less than additive effect would be expected and, in the extreme case, when the two doses were separated by a sufficient interval, the tumour incidence would be that induced by each dose separately, that is, at 80 weeks 5%, plus 5% of the animals in which cancer was not induced by the first dose.

Dimethylnitrosamine requires metabolism for its carcinogenic activity, so the possible effect of the first dose on the metabolism of the second must be considered. The second dose must give the same extent of reaction with the critical target, as happens when the single dose is increased from 16 mg kg^{-1} to 32 mg kg^{-1}. If the first dose of dimethylnitrosamine affects the metabolism of the second, the second dose required may be greater or less than 16 mg kg^{-1}. There is evidence that the extent of methylation produced in an organ by dimethylnitrosamine provides a measure of the extent of metabolism of dimethylnitrosamine in that organ, and of the extent of interaction of the ultimate carcinogen with the crucial cellular target[6]. By measuring the extent of methylation on the 7-position of guanine in the DNA of the kidneys produced by the second dose of dimethylnitrosamine, the second dose was adjusted so that, at every interval, the presumptive total extent of reaction with the

target cells was the same. The actual doses injected are given in the legend to Fig. 2.

The rats (in two batches) were given the first dose in April or May 1974. Of the 650 rats at the start of the experiment, 617 survived longer than 4 weeks. Of these, 437 are now dead, 259 of them with kidney tumours. The results (Fig. 2) seem to show that there is no repair (that is, the doses are at least cumulative) when the two doses are separated by no more than 4 d, but that apparently complete recovery takes place when the two doses are separated by 16 d or more. The reason for the considerable difference in incidence when the doses are separated by 1, 2 or 4 d is unknown. The steep dose–response curve magnifies small errors in dose to large differences in incidence, but the incidence when the doses were separated by 4 d is much greater than expected. The absolute amount of methylation of the N-7 and O^6 position of guanine, and the effect of the first dose on the excision of the O^6-methylguanine produced in the kidney DNA by the second dose, is being re-examined to try to explain it.

It is not likely that the recovery is the result of death of damaged cells and repopulation of the kidney by undamaged cells. Even 50 mg dimethylnitrosamine per kg body weight produces little necrosis in the kidney[7], and the half life of the N-7-methylguanine produced by the carcinogen in DNA is approximately the same in the kidneys of rats given 16 mg kg^{-1} as it is in the liver of rats given a dose which does not produce necrosis (ref. 8 and P. F. S. and R. Mace, unpublished). If there had been significant death and repopulation, the apparent half life of the alkylated base would have been shortened because the cells containing N-7-methylguanine would have been lost and replaced by cells which did not contain this alkylated base. The recovery therefore results not from repopulation but from repair of the critical damage which initiates the development of cancer.

The results reported here seem to be consistent with the hypothesis that the alkylation by dimethylnitrosamine and some alkylating agents of the O^6-position of guanine in DNA is the critical chemical lesion in the induction of cancer by these compounds[9] and that the animal can protect itself by excising this altered base from its DNA[10–13]. Although the rate and extent of the excision of dimethylnitrosamine-produced O^6-methylguanine from the kidney DNA of rats given exactly the diet and dose we used has not been measured, Nicoll *et al.*[12] found that after a dose of dimethylnitrosamine, which would be the equivalent of 13 mg kg^{-1} if the dietary conditions had been the same as ours, about 30% of the O^6-methylguanine was excised from the DNA between 6 h and 15 h after the dose. No further excision took place from 15 h to 48–72 h, after which excision resumed. The block of excision parallels, and might explain, the lack of recovery from the carcinogenic effect during the first 4 d (Fig. 2) but it remains to be seen whether, when the larger dose is given to rats on a protein-deficient diet, the rate and extent of excision follow more exactly the recovery from the carcinogenic effect.

Our conclusion that carcinogenic damage can be repaired differs from that of Druckrey[3], but the conclusion is not incompatible with his results, which were obtained from experiments in which carcinogen was administered repeatedly for long periods. It is not known what happens to the products of reaction between carcinogen and cell in these circumstances, but if the carcinogen acts by reacting with one of the bases in DNA so that its base-pairing properties are altered, the effect of this reaction product could be cumulative, either through accumulation of permanent base changes formed in the progeny cells during cell division, or through inhibition by the carcinogen of the excision from the DNA of the crucial reaction product. A large dose of dimethylnitrosamine inhibits the excision of O^6-methylguanine from rat liver DNA[14] and it is possible

that prolonged treatment with smaller amounts would have the same effect.

PETER F. SWANN
PETER N. MAGEE

Courtauld Institute of Biochemistry,
The Middlesex Hospital Medical School,
London W1P 5PN, UK

ULRICH MOHR
GERD REZNIK
UWE GREEN

Abteilung für experimentelle Pathologie,
Medizinische Hochschule Hannover,
3 Hannover 61,
Karl-Wiechert-Allee 9, FRG

DAVID G. KAUFMAN

Lung Cancer Group DCCP,
National Cancer Institute,
Bethesda, Maryland 20014

1 Cleaver, J. E., *Proc. natn. Acad. Sci. U.S.A.*, **63**, 428–435 (1969).
2 Miller, J. A., *Cancer Res.*, **30**, 559–576 (1970).
3 Druckrey, H., in *Potential Carcinogenic Hazards from Drugs* (edit. by Truhaut, R.), 60–78 (Springer, Berlin, 1967).
4 Nakahara, W., in *Hdb. allgemeinen Pathologie*, **6** (2) (edit. by Grundmann, E.), 595–650 (Springer, Berlin, 1975).
5 Elkind, M. M., and Sutton, H., *Radiat. Res.*, **13**, 556–593 (1960).
6 Swann, P. F., and McLean, A. E. M., *Biochem. J.*, **124**, 283–288 (1971).
7 Hard, G. C., and Butler, W. H., *Cancer Res.*, **30**, 2806–2815 (1970).
8 Margison, G. P., Capps, M. J., O'Connor, P. J., and Craig, A. W., *Chem-Biol. Interact.*, **6**, 119–124 (1973).
9 Lawley, P. D., *Mutat. Res.*, **23**, 283–295 (1974).
10 Goth, R., and Rajewsky, M. F., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 639–643 (1974).
11 Kleihues, P., and Margison, G. P., *J. natn. Cancer Inst.*, **53**, 1839–1841 (1974).
12 Nicoll, J. W., Swann, P. F., and Pegg, A. E., *Nature*, **254**, 261–262 (1975).
13 Margison, G. P., and Kleihues, P., *Biochem. J.*, **148**, 521–525 (1975).
14 Kleihues, P., and Margison, G. P., *Nature*, **259**, 153–155 (1976).

Changes in the behaviour of teratocarcinoma cells cultivated *in vitro*

MOUSE teratocarcinoma cells have been recognised as important material for studying *in vitro* the processes of determination and differentiation that take place during normal embryogenesis[1,2]. But in order for the system to be exploited maximally, it must be possible to manipulate at will the differentiation of mass cultures of pluripotent cells into endoderm, ectoderm and mesoderm derivatives. Martin and Evans showed that they could trigger the appearance of endoderm cells in undifferentiated cultures of the SIKR teratocarcinoma line by suspending aggregates of the cells in bacteriological Petri dishes to which they could not adhere[3,4]. Within about 2 d the aggregates developed into simple embryoid bodies, resembling egg cylinders of normal embryos with an outer layer of endoderm cells surrounding a core of undifferentiated cells. In some SIKR clones, but not others, the inner cells later differentiate into neuroectoderm and mesoderm and the embryoid bodies develop fluid cysts resembling the visceral yolk sac vesicles of normal early embryos cultured *in vitro*. Not all lines of pluripotent teratocarcinoma cells, however, can be induced to form embryoid bodies in such a convenient way (ref. 5 and my observations). This could be because certain teratocarcinomas are for some reason intrinsically unable to form embryoid bodies, or because the cells can lose this ability during cultivation *in vitro*. In the course of isolating a new clonal line of teratocarcinoma cells I noticed that their ability to differentiate was strongly influenced by the way in which they were maintained *in vitro*. As long as the cells were grown on layers of feeder fibroblasts they retained the ability to form embryoid bodies but this was progressively

lost as the cells were adapted to grow on gelatin-coated dishes.

The clone used for this work came from an OTT 6050 ascites tumour derived by Dr L. C. Stevens from a 6-d strain 129/Sv mouse embryo grafted to the testis of an adult[6]. A frozen stock of the ascites tumour was acquired by Dr M. J. Evans of University College, London and went through a further six passages *in vivo* before being cloned in this laboratory. Embryoid bodies in the ascites fluid were trypsinised and dispersed into 9-cm dishes containing 3×10^6 STO mouse fibroblasts that had been treated before plating out with mitomycin C (10 μg ml^{-1}) for 3.5 h (ref. 3). When colonies of undifferentiated embryonal cells had grown, the plates were trypsinised and single teratocarcinoma cells were isolated as described by Martin and Evans[3].

Cultures were maintained by plating approximately 5×10^5 cells into a 9-cm dish containing feeder fibroblasts in Dulbecco's modified Eagle's medium with 10% calf serum and 10% foetal calf serum (Flow Laboratories). The medium was changed daily and every 2–3 d the cells were detached with 0.05% trypsin and 0.6 mM EDTA in phosphate-buffered saline, pH 7.2, and subcultured. Until they were confluent, the cultures grew as homogeneous sheets of closely packed cells which spread evenly over the fibroblasts. Provided they were maintained in this way they kept their ability to differentiate into cystic embryoid bodies. The most efficient method of triggering differentiation was to suspend aggregates of the cells in bacteriological Petri dishes after a preliminary step to remove most of the feeder fibroblasts. The mixed cultures were trypsinised briefly to give clumps of 50–200 cells and these were plated out at low density on to a standard tissue culture dish in medium with 10% foetal calf serum. Overnight the fibroblasts attached tightly but the teratocarcinoma cells formed rounded aggregates which could be detached easily and transferred to bacteriological dishes. Two to three days after transfer, all aggregates resembled simple embryoid bodies: the outer endoderm cells had numerous vacuoles and microvilli, tight junctions and a well developed endoplasmic reticulum, and were separated from the inner core of undifferentiated cells by a thick basement membrane (analogous to the Reichert's membrane of normal embryos)[5]. By day 6 most of the bodies had developed a fluid-filled cyst and by 10 d about 20% of these contained small patches of haematopoietic cells.

Attempts to adapt the cells to grow in the absence of feeder fibroblasts did not succeed until (1) the dishes were coated with 1% gelatin (swine skin Type 1 300 bloom, from Sigma), (2) the cells were initially subcultured at high density and (3) the medium was supplemented with both 10% calf serum and 10% foetal calf serum. Even in these conditions, some cell death occurred during the first few subcultures, but eventually a morphologically homogeneous population of cells was obtained which grew rapidly (doubling time 20 h) in closely packed sheets and could be subcultured every 2 d at 1×10^6 cells per 9-cm dish. For the first few weeks these cells retained their ability to form cystic embryoid bodies when challenged by being put into bacteriological Petri dishes. With longer times in culture the cells would form only simple embryoid bodies. By 2 months (approximately 25 passages) only about 25% of the aggregates differentiated at all, whereas the remainder merely enlarged and became necrotic. By 4 months almost none of the clumps formed embryoid bodies after 5 d in suspension. These cultures were not entirely incapable of differentiating, however; if they were allowed to reach very high density on gelatin-coated dishes and the medium changed every day, scattered patches of endoderm cells appeared after about a week, followed by fibroblast-like cells and neurones, but this process was always accompanied by massive lysis of the undifferentiated cells. The loss of ability to differentiate rapidly seems to be irreversible

because it could not be restored by returning the deficient cells to feeder layers for several generations. The behaviour of these cells had then become very similar to that of other extensively studied OTT 6050-derived cell lines that were selected to grow in the absence of feeder fibroblasts on gelatin-coated[7] or untreated dishes[8].

These results suggest that during adaptation to growth on gelatin-coated dishes, strong selection occurs for cells which are better able to grow on a less adhesive substratum and perhaps because of this have a reduced probability of differentiating. Many of the advantages of using teratocarcinoma cells to study determination and differentiation are therefore lost if the cells are not maintained on feeder fibroblasts.

BRIGID L. M. HOGAN

Imperial Cancer Research Fund Laboratories,
Burtonhole Lane,
Mill Hill, London NW7 1AD, UK

1 Sherman, M. I., and Solter, D., (eds), *Teratomas and Differentiation* (Academic, New York, 1975).
2 Martin, G. R., *Cell*, **5**, 229–243 (1975).
3 Martin, G. R., and Evans, M. J., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 1441–1445 (1975).
4 Martin, G. R., and Evans, M. J., in *Teratomas and Differentiation* (edit. by Sherman, M. I., and Solter, D.), 169–187 (Academic, New York, 1975).
5 Sherman, M. I., in *Teratomas and Differentiation* (edit. by Sherman, M. I., and Solter, D.), 189–205 (Academic, New York, 1975).
6 Stevens, L. C., *Devl Biol.*, **21**, 364–382 (1970).
7 Nicolas, J. F., Dubois, P., Jacob, M., Gaillard, J., and Jacob, F., *Ann. Microbiol. Inst. Pasteur*, **126A**, 3–22 (1975).
8 Lehman, J. M., Speers, W. C., Swartzendruber, D. E., and Pierce, G. B., *J. cell. Physiol.*, **87**, 13–28 (1974).

Specific binding of a somatomedin-like polypeptide in rat serum depends on growth hormone

SOMATOMEDINS are growth hormone-dependent polypeptides that have been proposed as the mediators of the peripheral actions of growth hormone on skeletal tissue[1]. Several polypeptides with somatomedin activity *in vitro*, including somatomedins A, B and C and non-suppressible insulin-like activity soluble in acid ethanol (NSILA-s)[2-4], have been isolated from human plasma. Unique among polypeptide hormones, the somatomedins circulate in human blood complexed to binding proteins of larger molecular weight[2,4-7]. We report here the specific binding of a somatomedin-like polypeptide to distinct protein fractions in rat serum and further demonstrate that the pattern and level of this binding is dependent on growth hormone. We postulate that specific binding of somatomedin modulates the peripheral effects of somatomedin activity.

Results from our laboratory[8] have suggested that somatomedin-binding proteins prolong the half life of somatomedin activity in the circulation. In those studies, the half life of somatomedin activity derived from the serum of growth hormone-treated hypophysectomised (Hx) rats and injected into recipient Hx rats was determined. The half life of somatomedin activity from Hx animals treated with growth hormone for 10 d approached the half life of somatomedin activity derived from normal rats (3–4 h) and was significantly longer than the half life (< 15 min) of somatomedin activity derived from Hx rats treated with growth hormone for only 24 h. Furthermore, the half life of somatomedin activity that had been dissociated from serum proteins of large molecular weight by boiling in acid conditions or by chromatography on Sephadex G-50 in acid conditions before injection into Hx animals was less than 30 min. If this low molecular weight somatomedin activity was recombined with the serum proteins of high molecular weight from which somatomedin activity had been dissociated, the half life of this recombined somatomedin activity was prolonged to approximately 2 h. These experiments suggested that (1) the half life of somatomedin

Fig. 1 Sephadex G-200 elution pattern of different rat sera incubated with ^{125}I-MSA. MSA was purified from the conditioned media of a line of rat liver cells in culture by a modification of previous methods[9,10]. Briefly, MSA that had been eluted from Dowex was chromatographed on Sephadex G-75 in 1 N acetic acid. MSA eluted in three regions of activity as measured by stimulation of ^{3}H-thymidine incorporation into the DNA of chick embryo fibroblasts. Fractions constituting the middle region of activity were pooled and further purified by preparative disc electrophoresis at pH 2.7 in 9 M urea. The MSA preparation used yielded a single band in two different acrylamide gel electrophoresis systems. This MSA was approximately twice as active on a weight basis as porcine insulin in stimulating ^{3}H-thymidine incorporation into chick embryo fibroblast DNA. (This is the same specific activity reported by Smith and Temin relative to the same lot of porcine insulin[14].) MSA was iodinated by a modification of the chloramine T method[11]. A sample of 6 ng of ^{125}I-MSA ($150 \, \mathrm{Ci \, g^{-1}}$) was incubated with 0.5 ml of rat serum and 0.5 ml of Ca^{2+}-Mg^{2+}-free phosphate-buffered saline (Dulbecco's PBS) for 3 h at 30 °C. The incubation mixture was then applied to a 180-ml Sephadex G-200 column (2.6×40 cm) equilibrated and eluted with Dulbecco's PBS at 4 °C. Fractions of 2.2 ml were collected. Both ^{125}I-MSA radioactivity (●) and protein content (A_{280}) are plotted against fraction number and percentage of total bed volume. The excluded volume (V_0) was marked with blue dextran, and the included volume (indicated by the arrow) was marked with dinitrophenylalanine. The main protein peaks (— — — —) from left to right correspond to macroglobulins, gammaglobulins, and albumin. *a*, Elution pattern of normal rat serum incubated with ^{125}I-MSA. Peaks of ^{125}I-MSA are denoted by Roman numerals in order of decreasing molecular size. *b*, Elution pattern of serum obtained from Hx male Sprague–Dawley rats at least 10 d after hypophysectomy (performed when they were 90–100 g body weight)[8]. *c*, Elution pattern of serum obtained from Hx rats after 10 d of twice daily intraperitoneal injections of 100 μg of bovine growth hormone (chronic GH)[8]. *d*, Elution pattern of serum obtained from Hx rats after intraperitoneal injection of 100 μg of bovine growth hormone every 8 h for 24 h (acute GH)[8].

activity in the circulation was determined by its binding to large serum proteins and (2) these binding proteins were regulated by growth hormone.

To investigate directly the binding of a somatomedin-like polypeptide to rat serum, we have used multiplication stimulating activity (MSA), a 10,000-dalton polypeptide purified from the conditioned media of a line of normal rat liver cells in culture[9,10]. MSA has physicochemical similarities to the somatomedins purified from human plasma and shares with these polypeptides the ability to interact with the same receptors in various tissues and to stimulate

growth-related processes of certain cells and tissues in culture[2,9–13]. Specifically, the MSA preparation used for the studies reported here competes for somatomedin A binding to placental membranes[12]. Conversely, somatomedin A competes for the binding of this MSA to chick embryo fibroblasts, human fibroblasts and rat liver membranes[12]. An MSA preparation of similar specific activity also stimulates sulphate incorporation into chick embryo sternal cartilage at a concentration of $25 \, \mathrm{ng \, ml^{-1}}$ (unpublished experiments of K. Gibson and E. R. Froesch) and competes for NSILA-s binding to chick embryo fibro-
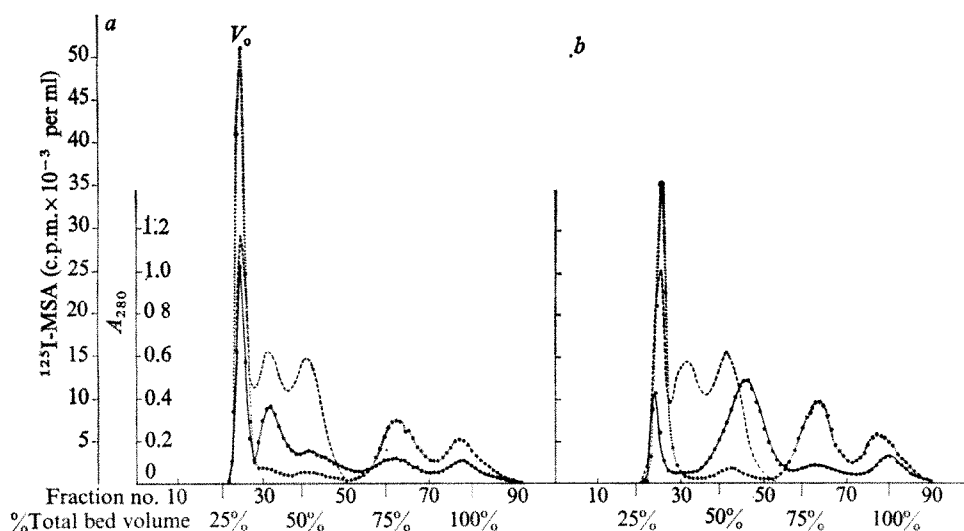
Fig. 2 Competition for ^{125}I-MSA binding to rat serum by unlabelled MSA. Normal rat serum (*a*) or Hx rat serum (*b*) were incubated for 3 h at 30 °C with 6 ng of ^{125}I-MSA in the presence (○--○) or absence (●———●) of 10 μg of unlabelled MSA as in Fig. 1. The samples were applied to a column of Sephadex G-200 and analysed as in Fig. 1. Protein content (— — — —) and radioactivity of the eluted fractions are plotted.

blasts (unpublished experiments of J. Zapf and E. R. Froesch).

^{125}I-MSA was incubated with normal rat serum and chromatographed on Sephadex G-200. Radioactivity appeared in five peaks, I–V (Fig. 1a). Only the binding to peaks II and III was specific, that is, it was abolished in the presence of excess (10 μg ml^{-1}) unlabelled MSA (Fig. 2). Binding of ^{125}I-MSA to peak I was nonspecific (Fig. 2), although the possibility of some specific binding of MSA to proteins in the peak I region has not been excluded completely. Peak IV represented free ^{125}I-MSA and peak V represented free ^{125}I as determined by the elution pattern of carrier free Na[125]I.
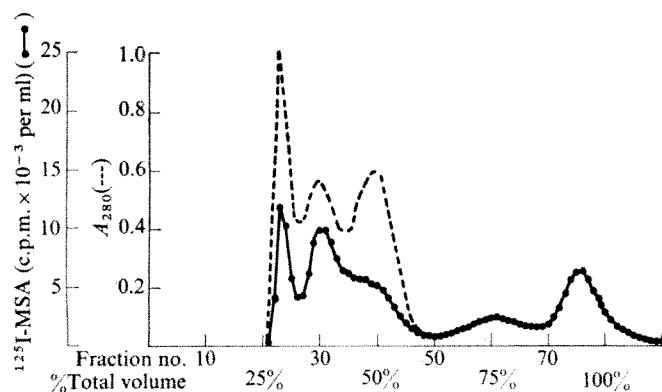
Fig. 3 MSA binding to a mixture of equal volumes of normal and Hx rat serum. ^{125}I-MSA (6 ng) was incubated with 0.25 ml of Hx rat serum, 0.25 ml of normal rat serum and 0.5 ml of Dulbecco's PBS for 3 h at 30 °C. The incubation mixture was then chromatographed on the same column of Sephadex G-200 as in Figs 1 and 2. Fractions were analysed for [125]I (●) and protein (– – – –).

The pattern of MSA binding to normal and Hx rat serum differed strikingly. Peak II, which represented the major specific binding area for MSA in normal rat serum, was absent in Hx rat serum. In this serum, ^{125}I-MSA bound specifically only to the peak III region (Figs 1b and 2b). MSA also bound to peak III in normal rat serum but to a lesser extent than in Hx rat serum (Fig. 1a and b).

When peaks II and III that had been isolated respectively from chromatography on Sephadex G-200 of normal and Hx rat serum were rerun individually over the same column, they eluted in a position identical to their position in whole serum (data not shown). This experiment excluded the conversion of one peak to the other. Furthermore, on the basis of the chromatographic pattern of a mixture of equal volumes of normal and Hx rat serum that had been incubated with ^{125}I-MSA, Hx rat serum did not inhibit binding of peak II (Fig. 3). Thus, an inhibitory component in Hx rat serum could not account for the absence of peak II MSA in Hx serum. These results indicated that MSA binds to distinct species, presumably proteins, in normal and Hx sera.

The different profiles of MSA binding to normal and Hx rat serum suggested that the binding pattern depended on growth hormone. This was confirmed by examining sera from Hx rats treated with growth hormone for different periods (Fig. 1c and d). Peak II, absent from Hx rat serum, reappeared after only 24 h of growth hormone treatment of Hx rats (Fig. 1d). After 10 d of growth hormone treatment, there was a further increase in the binding of ^{125}I-MSA to peak II that resulted in a binding profile virtually identical to that obtained with normal rat serum (Fig. 1c). Thus, although growth hormone increases the level of peak II MSA binding in Hx rats, the mechanism by which growth hormone enhances peak II binding remains undefined.

Growth hormone could induce the synthesis or release of the peak II-binding protein or it could modify a pre-existing protein such as peak III. Together with the observations of Cohen and Nissley[8], the above experiments suggest that the peak II binding protein prolongs the half life of somatomedin activity in serum. That is, serum containing the peak II MSA-binding protein yields a longer half life of somatomedin activity when injected into Hx animals than does serum without this binding protein.

The difference between MSA binding in normal and Hx rat serum was investigated further by a competitive binding assay (Fig. 4). After incubation of ^{125}I-MSA with unfractionated rat serum, unbound ^{125}I-MSA was separated from ^{125}I-MSA bound to serum proteins by adsorption to activated charcoal[5]. The binding of MSA to serum proteins was highly specific. Unlabelled MSA inhibited ^{125}I-MSA-binding to normal rat serum. Neither MSA that had been biologically inactivated by reduction and alkylation, human growth hormone, glucagon, human placental lactogen nor insulin at high concentrations competed effectively for ^{125}I-MSA binding (data not shown).

The specific binding of ^{125}I-MSA to normal rat serum was more than twice that to Hx rat serum (Fig. 4). In addition, serum from Hx rats treated with growth hormone for 10 d
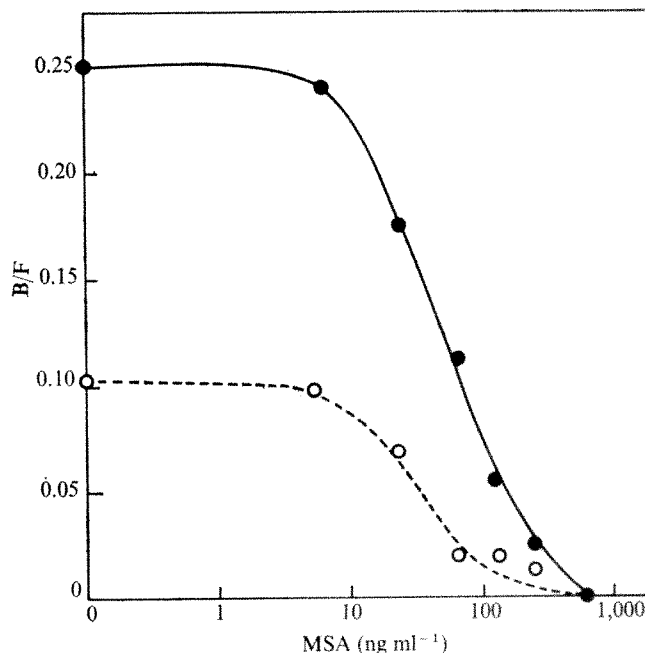
Fig. 4 Whole serum competitive binding assay. Normal (●) and Hx (○) rat sera were dialysed for 12 h at 4 °C against 250 volumes of Dulbecco's PBS and diluted with 3 volumes of the same buffer. Dialysed, diluted serum (25 μl) was incubated with 200 pg of ^{125}I-MSA in Dulbecco's PBS containing 0.2% human serum albumin in the presence or absence of unlabelled MSA at 30 °C in a total of 0.4 ml. After 3 h of incubation, 0.5 ml of Dulbecco's PBS containing activated charcoal in 2% human serum albumin was added to the reaction mixture[6]. The samples were equilibrated for 30 min at 0 °C and centrifuged at 2,500g for 10 min. The ^{125}I radioactivity in both supernatant liquid and the charcoal pellet were determined. (In a separate experiment, free[125]I-MSA quantitatively adsorbed to the activated charcoal. ^{125}I-MSA bound to serum proteins was not adsorbed.) The[125]I-MSA bound to serum proteins in the presence of unlabelled MSA (1 μg ml^{-1}) (nonspecific binding) was equal in both normal and Hx serum and has been subtracted from the total MSA before calculating the bound MSA to free MSA ratio (B/F). Nonspecific binding, equalled 24% of total binding for this experiment. B/F is plotted against increasing log concentrations of unlabelled MSA added to the reaction mixture. The binding of MSA to normal and growth hormone-treated Hx serum would underestimate specific MSA binding to these sera if endogenous somatomedins compete with [125]I- MSA for the binding protein(s).

bound 1.5 times the ^{125}I-MSA bound by serum from un-treated Hx rats (data not shown). The concentration of unlabelled MSA required to displace 50% of the specifically bound ^{125}I-MSA was approximately the same in normal, Hx and growth hormone-treated Hx serum (Fig. 4). These data suggest that normal and Hx rat serum have similar binding affinities but different binding capacities. This interpretation is supported by the predominant binding of MSA to peak II in the mixing experiment (Fig. 3). We presume that the difference in specific MSA binding by unfractionated normal and Hx serum reflects differences in the levels of peak II and peak III-specific binding proteins in these sera.

In summary, we have shown by Sephadex chromato-graphy and a competitive binding asay that normal rat serum specifically binds a somatomedin-like polypeptide; that this binding (peak II) is growth hormone dependent, and that its presence correlates with a prolonged half life of circulating somatomedin activity. The half life of somato-medin activity in Hx rat serum is short in spite of the presence of a specific somatomedin-binding protein (peak III). This could be explained by reduced binding capacity in Hx serum (Fig. 4) or by differences in the clearance or degradation of peak III or its complex with somatomedin.

Whatever the mechanism by which growth hormone alters somatomedin binding in serum, the dependence of somatomedin binding and half life on growth hormone adds an important 'fine-tuning' modulation to the control of growth by growth hormone. Indeed, the levels of a somato-medin-binding protein may determine the levels of somato-medin in the circulation. The absence of the peak II somatomedin-binding protein in Hx rats suggests that it may be difficult to produce *in vivo* growth-promoting effects with purified somatomedins due to the predicted rapid dis-appearance of these somatomedins from the circulation of the recipient Hx animal. Confirmation of this hypothesis awaits the purification of the somatomedin-binding protein and its injection into Hx animals in combination with somatomedin.

ALAN C. MOSES
S. PETER NISSLEY
KENNETH L. COHEN*

Endocrinology Section, Metabolism Branch,
National Cancer Institute,
National Institutes of Health,
Bethesda, Maryland 20014

MATTHEW M. RECHLER

Diabetes Branch,
National Institute of Arthritis, Metabolism and
Digestive Diseases,
National Institutes of Health,
Bethesda, Maryland 20014

*Present address: Department of Medicine, Yale University, New Haven, Connecticut 06520.

1 Daughaday, W. H., *et al.*, *Nature*, **236**, 107 (1972).
2 Van Wyk, J., *et al.*, *Recent Prog. Hormone Res.*, **30**, 259–318 (1974).
3 Froesch, E. R., *et al.*, *Adv. metabolic Dis.*, **8**, 203–209 (1975).
4 Hall, K., Takano, K., Enberg, G., and Fryklund, L., *Proc. int. Symp. Growth Hormone*, Milan, 1975 (Elsevier-Excerpta Medica, Amsterdam, in the press).
5 Zapf, J., Waldvogel, M., and Froesch, E. R., *Archs. Biochem. Biophys.*, **168**, 638–645 (1975).
6 Hintz, R. L., Orsini, E. M., and Van Camp, M. G., *The Endocrine Society 56th Annual Meeting Program*, A-71 (1974).
7 Yalow, R. S., Hall, K., and Luft, R., *Adv. metabolic Dis.*, **8**, 73–83, (1975).
8 Cohen, K. L., and Nissley, S. P., *Acta Endocr.* (in the press).
9 Dulak, N., and Temin, H., *J. Cell Physiol.*, **81**, 153–160 (1973).
10 Nissley, S. P., Passamani, J., and Short, P., *J. cell. Physiol.* (in the press).
11 Megyesi, K., *et al.*, *J. biol. Chem.*, **250**, 8990–8996 (1975).
12 Rechler, M. M., *et al.*, *The Endocrine Society 58th Annual Meeting Program*, 222 (1976).
13 Rechler, M. M., Podskalny, J. M., and Nissley, S. P., *Nature*, **159**, 134–136 (1976).
14 Smith, G. L., and Temin, H. M., *J. cell. Physiol.*, **84**, 181–192 (1974).

Serum can initiate DNA synthesis in cells rendered unresponsive to insulin and somatomedin

AN understanding of growth control in animal cells requires an appreciation not only of cellular events important to multiplication, but also of exogenous agents which either initiate or influence the 'decisions' of a cell to reproduce. Serum has been reported to contain factors required for cell movement[1], viability[2], ion transport[3] and control of cell density[4]. In addition, peptides from tissue or plasma have been identified (nerve growth factor (NGF)[5], somatomedin[6], non-suppressible insulin-like activity (NSILA)[7], multiplication-stimulating activity (MSA)[8], epidermal growth factor[9], fibroblast growth factor[10]) which seem to have a physiological role in the initiation of cell proliferation. Somatomedin, NSILA and MSA are probably similar, if not identical, substances and, like NGF, seem to be structurally related to insulin[11-13]. The ability of insulin to initiate DNA synthesis in various cells[14-17] may be due to cross reactivity of the hormone with receptor sites concerned with cell multiplication[12]. Recent reports suggest that serum contains additional substances that can influence the ability of a cell to respond to agents which initiate DNA synthesis[10,18,19]. While investigating the regulatory roles of insulin and serum in the initiation of DNA synthesis in quiescent chick embryo fibroblasts (CEF), we have found a distinctive difference in the initiation capacity of these agents after the inhibition of protein synthesis. We present here evidence for the existence of a regulatory component(s) in serum which recruits cells for growth from a stage in the cell cycle position (G_0?) which is re-fractory to mitogenic peptides.

As shown before[20-22], quiescent CEF initiate DNA syn-thesis 6-8 h after exposure to insulin or serum. Based on the incorporation of ^{3}H-thymidine into trichloroacetic acid (TCA)-insoluble material, the highest levels of DNA syn-thesis were initiated by insulin at a concentration of 1-5 μg ml^{-1} and approximated only 20-50% of the maximum levels stimulated by 3% serum (Fig. 1 legend). When serum-deprived CEF were treated with cycloheximide (5 μg ml^{-1}) before exposure to insulin, however, the subsequent insulin-induced DNA synthesis in washed, inhibitor-free cultures was found to be significantly reduced in comparison with previously uninhibited insulin-stimulated cultures. Figure 1 shows that the drug-induced inhibition of insulin-stimulated DNA synthesis was time dependent. Exposure to cyclohexi-mide for 6-8 h was required to reduce ^{3}H-thymidine in-corporation to 30-50% of that seen in uninhibited, hormone-stimulated cultures (Fig. 1). The percentage of insulin-treated cells with labelled nuclei was similarly decreased from 16% in uninhibited cells to 3% in inhibitor-free cells previously exposed to cycloheximide for 8 h (Fig. 2). In serum-deprived cultures, cycloheximide (5 μg ml^{-1}) inhibited incorporation of ^{3}H-amino acids into TCA-insoluble material by 70-80% as previously shown[24]. Immediately after the cells had been washed free of inhibitor, the incorporation of labelled amino acids returned to levels characteristic of serum-deprived control cultures (Fig. 1, inset). In contrast to the inhibition of insulin-stimulated DNA synthesis after exposure to cycloheximide, a similar cycloheximide treat-ment failed to reduce the serum-stimulated response (Figs 1 and 2). Similar results have been obtained with four different lots of calf serum. The fully recoverable serum-stimulated response has also been demonstrated with serum dialysed at a neutral pH (data not shown).

By measuring the DNA content of cultures, these results were further substantiated and extended to somatomedin, also an initiator of DNA synthesis in quiescent CEF[12]. In
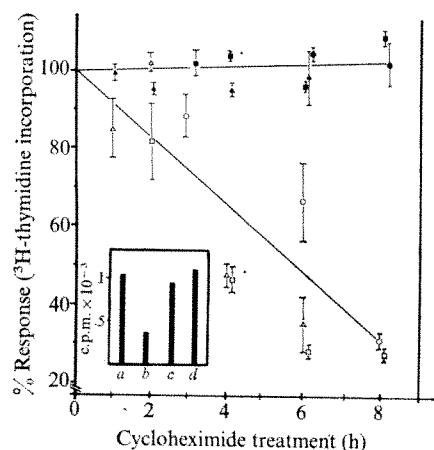
Fig. 1 Effect of cycloheximide pretreatment on insulin or serum-stimulated incorporation of ^{3}H-thymidine into TCA-insoluble material. CEF were prepared by trypsinisation of 11-d-old embryos[23]. Wells (diameter 16 mm) containing 12-mm glass coverslips were each seeded with 9×10^5 cells in 3% calf serum minimal essential medium (MEM) (1 ml per well). After incubation for 72 h (37 °C) in a humidified 5% CO_2 atmosphere, confluent cells (5×10^4–8×10^4 cells cm^{-2}) were washed once with Hanks' balanced salt solution (BSS) and further incubated in 1 ml of serum-free MEM for 24 h. Cells were then treated with cycloheximide at a final concentration of 5 µg ml^{-1} by the addition of small 10-µl samples from a concentrated stock of the drug. After a maximum cycloheximide treatment period of 8 h, all cells were washed twice with BSS, and then MEM with insulin (5 µg ml^{-1}) or 3% calf serum was added to wells in 1-ml volumes. After 6 h, 0.4 µCi of ^{3}H-thymidine (^{3}H-methyl, specific activity 64 Ci mmol^{-1}, Nuclear Dynamics) in MEM was added to each well. Twenty-four hours after the addition of hormone or serum, cells were washed once with cold saline (0.85%) and TCA-soluble material was extracted with cold 5% TCA (1 ml per well). After 10 min, the TCA was discarded and fresh 5% TCA added for a further 5 min. Coverslips were dipped in 95% ethanol and dried in counting vials at 50 °C. The TCA-insoluble radioactivity (c.p.m. per coverslip) was determined after the addition of Omnifluor–toluene scintillation fluid ($\triangle$, $\blacktriangle$; $\square$, $\blacksquare$). In one experiment, ($\bigcirc$, $\bullet$), after the second TCA wash coverslips were placed in small glass vials and the cells were digested with Protosol (New England Nuclear) for 2 h at 50 °C before addition of scintillation fluid. Percentage response is reported as the c.p.m. stimulated by insulin ($\triangle$, $\square$, $\bigcirc$) or serum ($\blacktriangle$, $\blacksquare$, $\bullet$) in three experiments after cycloheximide treatment for 1–8 h over uninhibited insulin-stimulated c.p.m. ($\triangle$, 25,820 c.p.m.$\pm$61; $\square$, 13,360 c.p.m.$\pm$960; $\bigcirc$, 39,095 c.p.m.$\pm$4174) or serum-stimulated c.p.m. ($\blacktriangle$, 54,414 c.p.m.$\pm$1537; $\blacksquare$, 50,160 c.p.m.$\pm$2606; $\bullet$, 84,662 c.p.m.$\pm$602), respectively. ^{3}H-thymidine incorporation in control cultures (MEM stimulated) for the three experiments averaged 2,545 c.p.m. The range and mean of duplicate determinations are shown. Inset, ^{3}H-amino acid incorporation into TCA-insoluble material in inhibited and uninhibited cultures. After treatment with (b, c) or without (a, d) cycloheximide for 6 h, cultures a and b were pulse labelled for 30 min at 37 °C with 2.5 µCi of ^{3}H-amino acids (Schwartz Mann, protein hydrolysate, 1 mCi ml^{-1}). c and d were washed twice with BSS and immediately pulsed with ^{3}H-labelled amino acids in fresh serum-free MEM for 30 min. Coverslips were then washed four times with cold saline and processed and counted as described above. The variation from the mean of duplicate samples was less than 10% in all cases. Amino acid incorporation was inhibited 71% in the presence of cycloheximide.

uninhibited cultures the DNA content 24 h after exposure to insulin, somatomedin (0.2 U ml^{-1}) or serum increased 18%, 20% and 50%, respectively (Fig. 3). In cells previously inhibited with cycloheximide (+CH) for 8 h, the increases in DNA initiated by insulin and somatomedin were either not observed or considerably reduced, again in contrast to the serum-initiated increases in DNA (Fig. 3). These data argue against the possibility that cycloheximide dissociates ^{3}H-thymidine transport and DNA synthesis, as shown with cytochalasin B treatment of 3T3 cells[26].

To explain the failure of the drug to diminish serum-stimulated DNA synthesis, it could be argued that the hormone-responsive cells inhibited by cycloheximide are

not part of the larger population of cells that responds to serum. Incorporation of TCA-insoluble ^{3}H-thymidine and autoradiography show however, that non-additive levels of DNA synthesis are obtained in cultures exposed to either of the hormones plus the maximum stimulatory concentration of serum (data not shown). This suggests that each of the hormones stimulates DNA synthesis in cell populations that also initiate DNA synthesis in response to serum. Similar findings have been reported before[7].
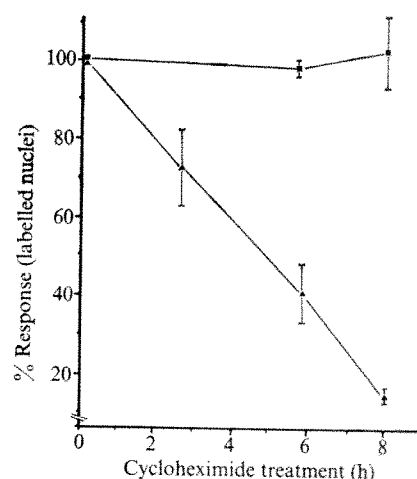
Fig. 2 Effect of cycloheximide pretreatment on insulin or serum-stimulated labelled nuclei. CEF were prepared as in Fig. 1. Wells (35 mm diameter) containing 22×22-mm glass coverslips were each seeded with 3×10^6 cells in 3% calf serum MEM (3 ml per well). After incubation for 72 h (37 °C) in a humidified 5% CO_2 atmosphere, confluent cells were washed once with BSS and further incubated in 2.5 ml of serum-free MEM. After 24 h in serum-free MEM, cells were treated with cycloheximide at a final concentration of 5 µg ml^{-1} by addition of 25-µl samples from a concentrated stock of the drug. After a maximum cycloheximide treatment period of 8 h, cells were washed twice with BSS, and then MEM with 5 µg ml^{-1} insulin or 3% calf serum was added to wells in 2.5-ml volumes. After 6 h 1 µCi of ^{3}H-thymidine in MEM was added to each well. Twenty-four hours after the addition of hormone or serum, cells were washed once with saline and processed for autoradiography as before[20]. Percentage response is reported as percentage of insulin- ($\blacktriangle$) or serum- ($\blacksquare$) stimulated cells with labelled nuclei after cycloheximide treatment for 3, 6 or 8 h over the percentage of uninhibited insulin-stimulated cells with labelled nuclei (16%$\pm$3) or serum-stimulated cells with labelled nuclei (58%$\pm$6), respectively. The percentage of cells with labelled nuclei after addition of MEM alone was 3%. The range and mean of duplicate determinations are shown.

The kinetics of the fully recoverable serum-stimulated DNA synthesis in CEF after cycloheximide treatment was monitored by pulse labelling cells with ^{3}H-thymidine at various times after addition of serum. Figure 4 shows that compared with uninhibited control cultures, both the onset and peak of serum-stimulated DNA synthesis were delayed by 3–5 h in cycloheximide-treated cells. In our hands, the timing of DNA synthesis in uninhibited insulin-stimulated CEF is similar to that of uninhibited serum-stimulated cells[20,27].

The delayed onset of DNA synthesis in inhibited cells after exposure to serum could account for the inability of the hormones to stimulate DNA synthesis, for during the additional time needed by cells to regain the capacity to synthesise DNA, the hormones could have been degraded. This seems unlikely, however, because media containing insulin 5 µg ml^{-1} remained fully active in stimulating DNA synthesis for at least 16 h after incubation with inhibited or uninhibited cultures. This latter observation (data not shown) was made by measuring ^{3}H-thymidine incorporation into TCA-insoluble material stimulated in fresh cell monolayers exposed to the 'used' insulin-containing media.

Thus, the ability of inhibited cells to synthesise DNA after
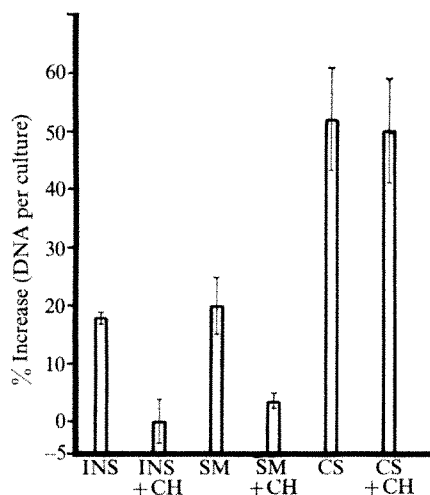
Fig. 3 Effect of cycloheximide pretreatment on the DNA content of CEF after the addition of insulin (INS), somatomedin (SM) or calf serum (CS). CEF were prepared as for Fig. 1. Plastic flasks (72 cm^2) were each seeded with 2.2×10^7 cells in 3% calf serum MEM (15 ml per flask). After incubation for 72 h (37 °C) in a humidified 5% CO_2 atmosphere, confluent cells were washed once with BSS and further incubated in 15 ml of serum-free MEM. After 24 h in serum-free MEM, cultures were treated with cycloheximide (5 µg ml^{-1}) and after 8 h were washed twice with 15 ml of BSS. Cells were exposed to MEM, INS (5 µg ml^{-1}) S, SM (0.2 µg ml^{-1}) or 3% CS. Twenty-four hours later the cells were washed once with saline, trypsinised (0.5% trypsin), pelleted by centrifugation and counted. The DNA content per flask (3×10^6–5×10^6 cells) was determined after treatment of cells with 1 N perchloric acid at 70 °C for 20 min, using the method of Burton[25]. The percentage increase in DNA content was determined using the A_{600} values obtained for MEM-stimulated cells as baseline (0% increase). For each bar, the variation and mean of two experiments is given. In each experiment, all points were determined in duplicate and averaged.

exposure to serum seems to require additional time-dependent cellular changes not required of uninhibited cells. In view of the inability of insulin or somatomedin to induce DNA synthesis in cycloheximide-treated cells, these further changes seem to be initiated by a serum component(s) distinct from those which signal DNA synthesis directly. In this regard, serum seems to be performing a more complex function than simply supplying inhibited cells with an exogenous nutrient needed for DNA synthesis. If this were the case, the timing of DNA synthesis might be expected to be unaffected by cycloheximide treatment[28].
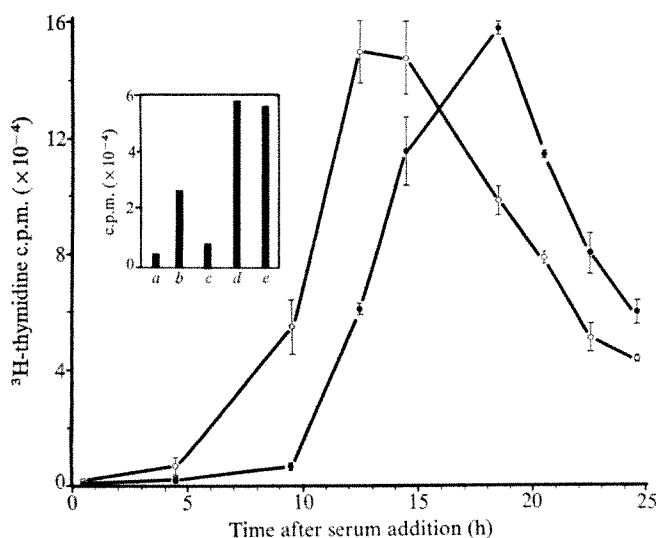
To summarise and further clarify our ideas, we suggest that the inhibition of protein synthesis results in the loss of some cellular protein(s) required directly or indirectly for the induction of DNA synthesis by insulin and somatomedin, and that these agents cannot reinitiate the synthesis of this protein. An additional regulatory component(s) in serum is proposed that can signal the synthesis of this missing cellular protein(s), thus permitting the subsequent induction of DNA synthesis by factors in serum functionally analogous to insulin and somatomedin. In terms of the eukaryotic cell cycle, this additional serum component(s) could have a role in the recruitment of cells from G_0 into G_1.

We thank J. J. Van Wyk for helpful suggestions and for insulin-free somatomedin (specific activity 8.7 U per mg protein). This research was aided by a grant from the American Cancer Society to the University of North Carolina. I.B.B. is the recipient of a USPHS Career Development Award from the NIAID.

JOHN GROELKE
JOEL B. BASEMAN

Department of Bacteriology and Immunology,
University of North Carolina, Chapel Hill,
North Carolina 27514

1 Lipton, A., Klinger, I., Paul, D., and Holley, R. W., *Proc. natn. Acad. Sci. U.S.A.*, **68**, 2799–2801 (1971).
2 Paul, D., Lipton, A., and Klinger, I., *Proc. natn. Acad. Sci. U.S.A.*, **68**, 645–648 (1971).
3 Cunningham, D. D., and Pardee, A. B., *Proc. natn. Acad. Sci. U.S.A.*, **64**, 1049–1056 (1969).
4 Holley, R. W., in *Control of Proliferation in Animal Cells*, (edit. by Clarkson, B., and Baserga, R.), 13–18 (Cold Spring Harbor Laboratory, New York, 1974).
5 Levi-Montalcini, R., and Angeletti, P., *Physiol. Rev.*, **48**, 534–569 (1968).
6 Van Wyk, J. J., and Underwood, L. E., *A. Rev. Med.*, **26**, 427–442 (1975).
7 Morrell, B., and Froesch, E. R., *Eur. clin. Invest.*, **3**, 119–123 (1973).
8 Dulak, N. C., and Temin, H. M., *J. cell. Physiol.*, **81**, 153–160 (1973).
9 Cohen, S., and Taylor, J. M., *Rec. Prog. Hormone Res.*, **30**, 533–550 (1974).
10 Gospodarowicz, D., *Nature*, **249**, 123–127 (1974).
11 Frazier, W. A., Angeletti, R., and Bradshaw, R. A., *Science*, **176**, 482–487 (1972).
12 Van Wyk, J. J., *et al.*, *Adv. metabolic Dis.*, **8**, 128–150 (1975).
13 Rechler, M. M., and Podskalny, J. M., *Nature*, **259**, 134–136 (1976).
14 Temin, H. M., *J. cell. Physiol.*, **69**, 377–384 (1967).
15 Lockwood, D. H., Voytovich, A. E., Stockdale, F. E., and Topper, Y. J., *Proc. natn. Acad. Sci. U.S.A.*, **58**, 658–664 (1967).
16 Griffiths, J. B., *Expl Cell Res.*, **75**, 47–56 (1972).
17 Paul, D., and Walter, S., *J. cell. Physiol.*, **85**, 113–124 (1975).
18 Leffert, H. K., *J. Cell Biol.*, **62**, 767–779 (1974).
19 Nadal, C., and Boffa, G. A., *Cell Tissue Kinet.*, **8**, 297–305 (1975).
20 Baseman, J. B., and Hayes, N. S., *J. Cell Biol.*, **67**, 492–498 (1975).
21 Vaheri, A., Ruoslahti, E., Hovi, T., and Nordling, S., *J. cell. Physiol.*, **81**, 355–364 (1973).
22 Temin, H., *J. cell. Physiol.*, **78**, 161–170 (1971).
23 Rubin, H., *Virology*, **4**, 533–562 (1957).
24 Willis, M. V., Baseman, J. B., and Amos, H., *Cell*, **3**, 179–184 (1974).
25 Burton, K., *Methods in Enzymology*, **12**, 163–166 (1968).
26 Brownstein, B., Rozengurt, E., Jimenez, De Asua, L., and Stoker, M., *J. cell. Physiol.*, **85**, 579–585 (1975).
27 Baseman, J. B., Paolini, D., and Amos, H., *J. Cell Biol.*, **60**, 54–64 (1974).
28 Boynton, A. L., Whitfield, J. F., and Isaacs, R. J., *In Vitro*, **12**, 120–123 (1976).

Fig. 4 Effect of cycloheximide pretreatment on the timing of DNA synthesis in CEF after the addition of serum. CEF were prepared as for Fig. 1. Glass bottles (11 cm^2) were each seeded with 4.5×10^6 cells in 3% calf serum MEM (3 ml per bottle). After incubation for 72 h in a humidified 5% CO_2 atmosphere, confluent cells were washed once with BSS and further incubated in 3 ml of serum-free MEM. After 24 h in serum-free MEM, cultures were treated with or without cycloheximide (5 µg ml^{-1}) and after 8 h were washed twice with 3 ml of BSS. Cells were treated with MEM containing 3% calf serum and pulse labelled with ^{3}H-thymidine (5 µCi per bottle) for 1-h intervals at the times indicated. Pulse-labelled cells, washed once with cold saline (3ml) were scraped into 1 ml of saline to which 7.5% TCA was added. TCA-insoluble material was collected by centrifugation, washed once with 5% TCA, and dissolved in 0.1 M NaOH (0.75 ml). Samples of 0.25 ml were added to 5 ml of scintillation fluid (Triton X-Omnifluor–toluene) and counted. The range and mean of duplicate determinations are shown. ○, Uninhibited cultures; ●, cycloheximide-pretreated cultures. Inset, DNA synthesis in control cultures for 24 h after the addition of hormone or serum. After treatment with (c, d) or without (a, b, e) cycloheximide for 8 h, cells were washed and exposed to a, MEM; b, c, insulin; d, e, serum as described above. ^{3}H-thymidine (0.4 µCi per bottle) was added 6 h after the addition of hormone or serum. Cells were collected 18 h later and the TCA-insoluble material was prepared and counted as described above. Insulin-stimulated DNA synthesis was inhibited 69% by cycloheximide pretreatment.

Magnesium required for serum-stimulation of growth in cultures of chick embryo fibroblasts

THE addition to culture medium of phosphorylated compounds which preferentially bind Mg^{2+} substantially reduces the rate of DNA synthesis and metabolism of chick embryo fibroblasts (CEF)[1]. This has suggested that Mg^{2+} might be an intracellular regulator for coordinate control of the rate of cell growth and metabolism in animal cells[1]. We report here that the omission of Mg^{2+} from the culture medium can have considerable effects on cell growth and macromolecular synthesis by CEF if the Mg^{2+} is withdrawn at the time cells are being stimulated to grow by means of an increase in the concentration of serum in the medium. These observations are discussed in terms of a model in which the intracellular concentration of free Mg^{2+} coordinately regulates the rate of metabolism and cell growth through its role as an essential cofactor for transphosphorylation reactions.

Primary cultures of CEF were prepared as before[2]. Secondary cultures were plated 3–5 d before use at 10^6 cells per 60-mm Falcon culture dish in medium 199, containing 2% tryptose phosphate broth (TPB) and 1% chicken serum. During experimental manipulation, medium 199 containing 2% TPB was again used at the indicated Mg^{2+} concentration, but with dialysed serum present. Experimental techniques for isotopic labelling, scintillation counting and autoradiography have been described before[3,4]. Concentrations of Mg^{2+} in the culture medium were determined with a Perkin–Elmer atomic absorption spectrophotometer.
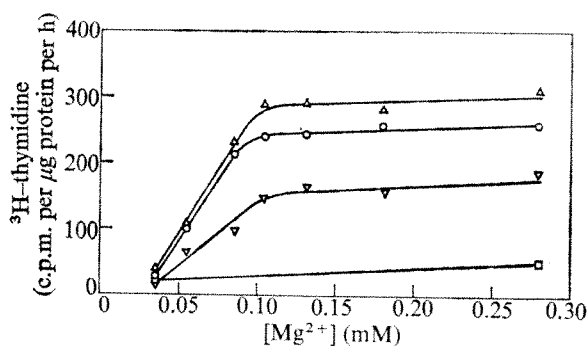
Fig. 1 ^{3}H-thymidine incorporation by CEF as function of the concentration of Mg^{2+} and calf serum. Four-day-old cultures were changed to fresh medium at the indicated concentrations of Mg^{2+} and calf serum; 20 h later the rate of DNA synthesis was measured by ^{3}H-thymidine (0.6 µCi ml^{-1} medium) incorporation into acid insoluble material for 60 min (calf serum: $\triangle$, 8%; $\bigcirc$, 4%; $\bigtriangledown$, 1%; $\square$, 0.0%).

Stimulation of cell growth by serum causes an increase in the rate of ^{3}H-thymidine incorporation by CEF that is proportional to the increase of the number of cells in the S phase of the cell cycle[5]. But as Fig. 1 shows, reducing the Mg^{2+} concentration of the medium from the standard 0.8 mM concentration to 0.04 mM prevents the increase of ^{3}H-thymidine incorporation. Below 0.1 mM Mg^{2+} the rate of ^{3}H-thymidine incorporation was directly proportional to the Mg^{2+} concentration of the medium at any concentration of calf serum (or chicken serum) from 1% to 8%.

There have been reports that the addition of Mg^{2+} to culture medium is not essential for the growth of mouse L cells[6] and HeLa cells[7]. We suppose that either of those cell lines have a slightly lower Mg^{2+} requirement for growth or, in the case of the report on mouse L cells[6], the medium
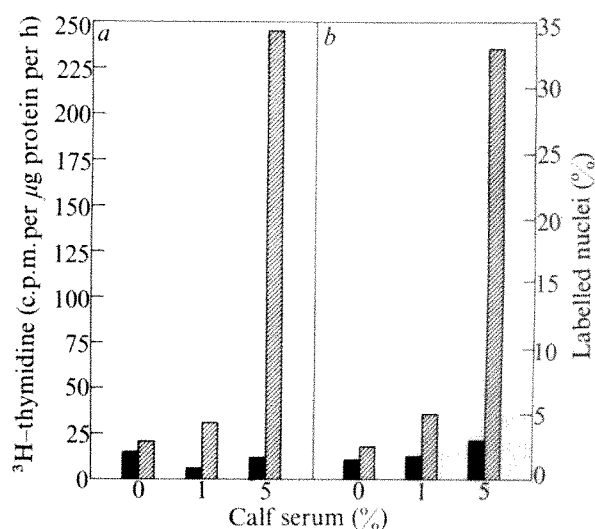
Fig. 2 Comparison of total ^{3}H-thymidine incorporation and autoradiography as measures of the rate of DNA synthesis. Four-day-old cultures were changed to medium containing 0.04 mM Mg^{2+} and the indicated concentrations of dialysed calf serum. After 20 h, half of the cultures were stimulated to grow by increasing the Mg^{2+} concentration of the media to 0.84 mM. Twenty hours later ^{3}H-thymidine (0.6 µCi ml^{-1}) was added to the medium for 60 min. *a*, Half of the cultures were used to determine total ^{3}H-thymidine incorporation into acid insoluble material; *b*, parallel cultures prepared for autoradiography as described before[4]. Solid columns, 0.04 mM Mg^{2+}; hatched columns, 0.84 mM Mg^{2+}.

used might have contained more Mg^{2+} as a contaminant than was present in ours. As both of these cell lines have malignant characteristics, a reduced requirement for Mg^{2+} could also be characteristic of malignant transformation.

Figure 2 shows that Mg^{2+} deprivation of cells reduced ^{3}H-thymidine incorporation mainly by decreasing the rate
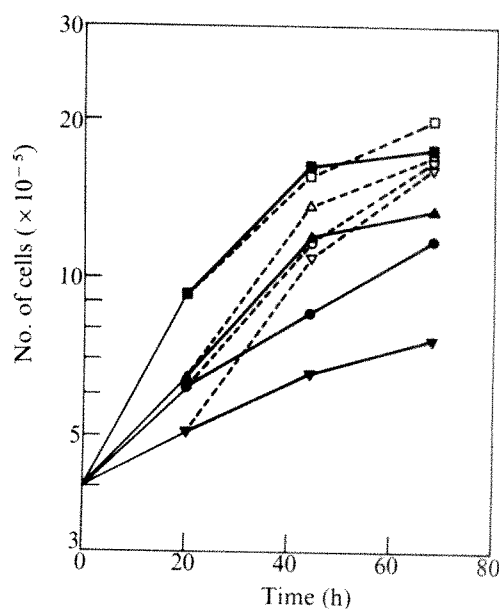
Fig. 3 Reversibility of inhibition of cell growth by Mg^{2+} deprivation. At zero time 5×10^5 CEF in medium 199 ($-Mg^{2+}$) containing 2% TPB, 5% dialysed calf serum and the indicated concentrations of Mg^{2+} were added to 60-mm culture plates. In these conditions, approximately 80% of the cells attach to the plates by 3 h. At 20 h half of all of the cultures were changed to medium containing the standard 0.84 mM Mg^{2+}. Then at 20, 44 and 68 h the number of cells on each plate was determined, after trypsin treatment, with a Coulter counter. Original medium: $\blacktriangledown$, 0.04 mM Mg^{2+}; $\bullet$, 0.06 mM Mg^{2+}; $\blacktriangle$, 0.11 mM Mg^{2+}; $\blacksquare$, 0.84 mM Mg^{2+}. The open symbols with dashed lines indicate the corresponding cultures changed at 20 h to medium containing 0.84 mM Mg^{2+}.

of entry of cells into the S phase of the cell cycle, not just by slowing the rate of DNA synthesis in individual cells. When Mg^{2+} was added back to the culture medium of cells that had been deprived of Mg^{2+} for 24 h, there was stimulation of both ^{3}H-thymidine incorporation and percentage ^{3}H-thymidine-labelled nuclei : this was directly related to the serum concentration in which the cells had been maintained during this period. In the case of Mg^{2+}-deprived cells maintained in medium containing 5% calf serum there was a 17-fold increase in the rate of ^{3}H-thymidine incorporation 20 h after addition of Mg^{2+} to the culture medium. Autoradiograms of parallel cultures revealed a 10-fold increase in the number of ^{3}H-thymidine-labelled nuclei at this time. If ^{3}H-thymidine incorporation accurately reflected the rate of DNA synthesis, then after the addition of Mg^{2+} individual cells made DNA only 1.7 times faster, but 10 times more cells were in S phase. Likewise with Mg^{2+}-deprived cultures maintained in 1% calf serum there was a fivefold increase in ^{3}H-thymidine incorporation and a threefold increase in the percentage of labelled nuclei 20 h after addition of Mg^{2+}.

Figure 3 shows the inhibition of cell multiplication by Mg^{2+} deprivation, and its reversal when Mg^{2+} was restored. At zero time 5×10^5 CEF in medium containing the indicated concentrations of Mg^{2+} were plated on 60-mm culture dishes. About 80% of the cells became attached at all Mg^{2+} concentrations. At 20 h half of all the cultures were moved to medium containing 0.84 mM Mg^{2+}. Then at 20, 44 and 68 h the cells on each plate were counted, after trypsin treatment, with a Coulter counter. By 20 h after plating-out of cultures, the cells growing in medium containing 0.84 mM Mg^{2+} had doubled in number and continued

Fig. 4 Effect of Mg^{2+} deprivation on ^{3}H-proline, ^{3}H-uridine and ^{3}H-thymidine incorporation. Five-day-old cultures were changed to medium containing 5% dialysed calf serum and either 0.84 mM Mg^{2+} or 0.04 mM Mg^{2+}. At various times the rates of synthesis of protein, RNA and DNA were determined by incorporation of: *a*, ^{3}H-proline (1.5 µCi ml^{-1}); *b*, ^{3}H-uridine (0.1 µCi ml^{-1}); or *c*, ^{3}H-thymidine (0.6 µCi ml^{-1}), respectively, into acid-insoluble material during 60-min incubations. ○, 0.04 mM Mg^{2+} ; ●, 0.84 mM Mg^{2+}.

to multiply until 48 h. In cultures with 0.11 mM Mg^{2+} or less in the growth medium, however, the growth rate showed a dependence on Mg^{2+} concentration. Mg^{2+} deprivation did not reduce the growth rate of these cells by simply decreasing their viability. Addition of 0.8 mM Mg^{2+} to the growth medium restored the increase of these cells to the maximum rate and allowed them to achieve the population density of cells in control cultures. This illustrates the possible potential of alterations in intracellular Mg^{2+} for the physiological regulation of cell growth.

As well as preventing serum-stimulation of cell growth, Mg^{2+} deprivation resulted in substantial decreases in the rates of synthesis of RNA and protein (Fig. 4). Twenty hours after the Mg^{2+} concentration had been decreased to 0.04 mM the rates of incorporation of ^{3}H-proline and ^{3}H-uridine were reduced 3.5-fold and 10-fold, respectively, from the level of control cells growing in 5% calf serum. Therefore Mg^{2+} deprivation seems not only to inhibit the rate of growth of CEF but also to inhibit their rate of macromolecular synthesis and presumably metabolism as a whole. This coordinate metabolic response of cells when the extracellular concentration of Mg^{2+} is reduced is also displayed on removal of serum or contact inhibition of growth[4]. Many agents that stimulate growth, such as serum, insulin and pH manipulation of the medium, seem to increase cellular functions in a coordinate manner[5]. We believe that the intracellular concentration of free Mg^{2+} has a central role in this coordinate metabolic response of animal cells to all these agents. Other divalent cations including Zn^{2+} (ref. 9) and Ca^{2+} (ref. 10) may have subsidiary roles.

In previous experiments, in which $Na_4P_2O_7$ or adenine nucleotides have been used to complex Mg^{2+} in the medium, similar effects on cellular metabolism have been noted and a model for coordinate control of cell metabolism has been proposed[1]. According to this model, intracellular Mg^{2+} is limiting in the formation of adenine nucleotide–Mg^{2+} complexes, essential substrates for transphosphorylation reactions. Transphosphorylation reactions, in turn, are viewed as the regulatory steps in metabolic pathways. This seems to be true in the glycolytic pathway, the one example where all the substrates and products of the pathway are readily assayed, since a selective activation of transphosphorylation reactions takes place when rates of glycolysis are increased[11,12]. The control of the intracellular level of free Mg^{2+} is brought about, in the context of this model, by changes in the competitive binding capacities of membranes and adenine nucleotides for Mg^{2+}. Because most of the cellular Mg^{2+} are presumably not free inside the cell but bound to membranes and other cellular constituents[13,14], serum and other agents that stimulate cell growth are postulated to alter membrane structure so as to release small amounts of membrane-bound Mg^{2+}. Chelation of Ca^{2+} in the culture medium by EGTA has been reported to inhibit cell growth in chick fibroblasts[10]: this might be the result of a competitive binding of Mg^{2+} and Ca^{2+} to membranes. The removal of sufficient Ca^{2+} from cellular membranes could result in a replacement by Mg^{2+} (ref. 13), resulting in a lower intracellular concentration of free Mg^{2+}. Indeed, it has been shown that the removal of Ca^{2+} from the medium markedly increases the sensitivity of cells to Mg^{2+} deprivation[15].

JAMES KAMINE
HARRY RUBIN

*Department of Molecular Biology
and Virus Laboratory,
University of California,
Berkeley, California 94720*

1 Rubin, H., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 3551–3555 (1975).
2 Rein, A., and Rubin, H., *Expl cell Res,*. **49**, 666–678 (1968).
3 Gurney, T., *Proc. natn. Acad. Sci. U.S.A.*, **62**, 906–911 (1969).
4 Rubin, H., and Koide, T., *J. Cell Biol.*, **56**, 777–786 (1973).
5 Rubin, H., and Steiner, R., *J. Cell. Physiol.*, **85**, 261–270 (1975).
6 Yang, D., and Morton, H. J., *In Vitro*, **6**, Abstract 54 (1970).
7 Eagle, H., *Archs Biochem. Biophys.*, **61**, 356–366 (1956).
8 Rubin, H., and Koide, T., *J. Cell. Physiol.*, **86**, 47–58 (1975).
9 Rubin, H., *Proc. natn. Acad. Sci. U.S.A.*, **69**, 712–716 (1972).
10 Balk, S. D., Whitfield, J. F., Youdale, T., and Braun, A. C., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 675–679 (1973).
11 Fodge, D., and Rubin, H., *Nature new Biol.*, **246**, 181–183 (1973).
12 Singh, V., Singh, M., August, J., and Horecker, B., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 4129–4132 (1974).
13 Sanui, H., *J. Cell. Physiol.*, **75**, 361–368 (1970).
14 Rose, I., *Proc. natn. Acad. Sci. U.S.A.*, **6**, 1079–1086 (1968).
15 Rubin, H., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 168–172 (1976).

C1q production and secretion by fibroblasts

THE serum protein C1q is a subcomponent of the first component of the complement system. Combination of C1q with antigen–antibody complexes activates the other subcomponents of C1 (C1r and C1s) and sets in motion a series of events which results in activation of the entire complement system. The primary structure of C1q is unusual in that 40% of its amino acid sequence is composed of repeating X-Y-Gly triplets where X is often proline and Y often hydroxyproline or hydroxylysine, indicating a striking similarity to collagen[1,2]. It has also been proposed that these collagen polypeptide-like stretches in C1q associate in triple helical fashion[3]. Because of the similarities between these portions of C1q and collagen the present investigation was undertaken to determine whether C1q is synthesised by collagen-producing cells, namely fibroblasts.

Three fibroblast lines were used: human lung fibroblasts (from 13-week embryos) obtained from Gibco-Biocult (Paisley, Scotland); human skin fibroblasts (from 14–16-week embryos) and rat skin fibroblasts (from 16–18-day embryos) cultured in this laboratory. All cell lines were grown in air on glass slides in Petri dishes, or in 50-cm² glass culture flasks in HEPES buffered minimum essential medium containing 10% foetal calf serum, 2 mM glutamine, 0.25 mM ascorbate, 0.025 μM ferric nitrate and 50 units of penicillin and streptomycin ml^{-1}. All lines formed collagen fibres (demonstrated by reticulin staining) and contained prolyl hydroxylase. The cells used in all experiments had been subcultured four to six times.

C1q was localised in all three cell lines growing on glass slides, by an immunoperoxidase procedure[4] using a monospecific rabbit antiserum to human C1q (Hoechst Hounslow, England) and to rat C1q. As shown in Fig. 1, fibroblasts bind antibody to C1q avidly. This reaction (Fig. 1) is specific

Fig. 1 Localisation of C1q in cultured human fibroblasts. The cells were treated with monospecific antibody to human C1q and goat anti-rabbit IgG peroxidase conjugate by an immunoperoxidase procedure[5].
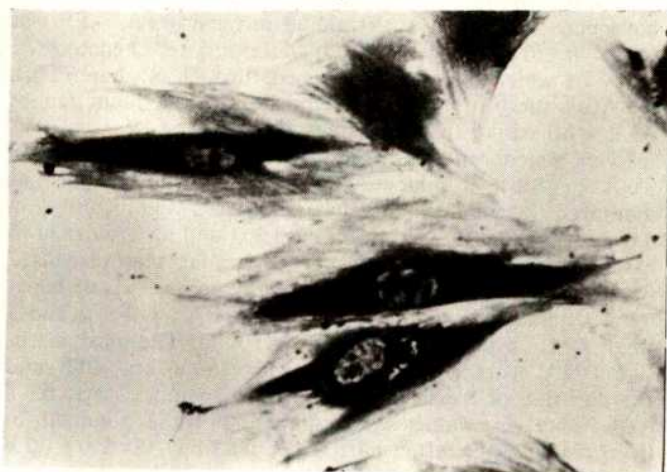
and indicates that C1q was present in the fibroblast cytoplasm of all cell lines as shown by the following control experiments. Fibroblasts overlaid by normal rabbit serum, or antiserum to C1q absorbed with pure C1q (ref. 5) or with fresh human or rat serum before processing through the immunohistochemical procedure failed to stain any of the fibroblast lines (Fig. 2). Antisera to C1q absorbed with heat-denatured normal serum (which renders C1q incapable of reacting with its antibody) did not abolish the reaction seen in Fig. 1. To exclude the possibility that antiserum to C1q cross reacts with collagen two further absorption experiments were done: previous absorption of antiserum to human C1q with neutral salt-soluble collagen failed to alter the reaction seen in Fig. 1; likewise antiserum to human C1q did not react in agar diffusion plates with a mixture of Type I and Type III procollagen purified from human embryonic skin[6]. On this evidence, therefore, it seems unlikely that anti-C1q was reacting with intracellular collagen antigen.
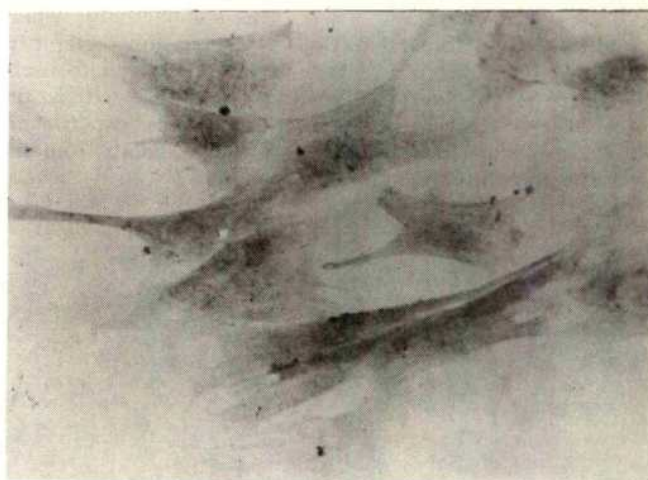
Fig. 2 Cells treated as in Fig. 1 except that normal rabbit serum was used instead of the antibody. No staining can be seen except that due to osmium tetroxide.

To determine whether C1q was also secreted by these cells, human fibroblasts from embryonic skin were grown to mid log phase under the same conditions in 50-cm² flasks. Each culture was then washed wtih 0.15 ml NaCl (2 × 5 ml) and incubated in 10 ml of serum-free minimum essential medium (supplemented as above) containing 2.4 μCi ml^{-1} of 3,4-^{3}H-proline (specific activity 20 mCi mmol^{-1}) for 60 min at 37 °C. The cell layers were then washed with 7 ml of 0.15 M NaCl and the medium and NaCl wash pooled and frozen at −20 °C until used. The cells were collected in 0.15 M NaCl, resuspended to 10^7 cells ml^{-1} in 0.1 M phosphate buffer (pH 7.2) containing 0.15 M NaCl and 0.01 M

Table 1 Immunoprecipitation of C1q from human fibroblast cultures and medium

Experiment	Aliquot (ml)	C1q precipitated per 10^7 cells* (c.p.m.)
Cell layer	0.15	1,407
Medium	2.5	4,225
	5.0	3,600

Human fibroblasts were cultured, pulsed with ^{3}H-proline, collected, sonicated and centrifuged as described in the text. Labelled C1q was precipitated from 0.15 ml aliquots of the cell supernatants by the addition of 0.1 ml of fresh human plasma (as carrier C1q) and rabbit antiserum to C1q. Labelled C1q was precipitated from aliquots of the medium in a similar way (see text).
*The number of counts in the control tubes (with normal rabbit serum) was 20% and 5% of that found in those containing anti-C1q for the cell layer and medium respectively. All results are the average of duplicate estimations.

EDTA, sonicated and centrifuged at 100,000g for 90 min. Fresh normal human serum (as carrier C1q) was added to the supernatant and to medium dialysed exhaustively against phosphate-buffered saline containing 0.01 M EDTA. Rabbit anti-human C1q was then added in threefold excess to the cell supernatants and medium, incubated at 37 °C for 30 min and overnight at 4 °C. In control tubes normal rabbit serum was added in the same proportion as the rabbit antiserum to C1q. The immunoprecipitates were collected, washed, dissolved in 1 N NaOH, and the radioactivity in each precipitate counted in Bray's solution. As shown in Table 1, C1q is present not only in the cells themselves but is also rapidly secreted by them such that in a 1-h pulse period approximately three times more labelled C1q is present in the medium than in the cell layer. It should be noted that the total amount of protein present in the immune precipitate was constant. For example, in the quantitation of C1q in medium (Table 1) the total amount of protein in each precipitate was 0.39 mg and 0.33 mg using 2.5 and 5 ml of medium respectively. This, therefore, excludes the possibility that the radioactivity which coprecipitated with carrier C1q was due to nonspecific precipitation or absorption of tritium to the precipitate itself. These data also argue against the possibility that the reaction product seen in immunohistochemical experiments (Fig. 1) was due to absorption of preformed C1q from the serum containing medium by fibroblasts.
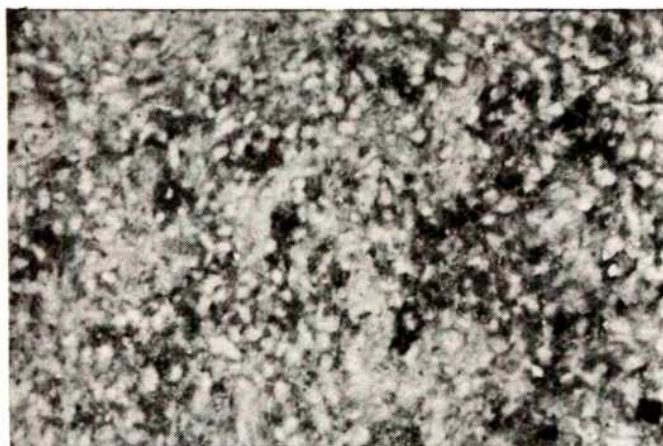
Fig. 3 C1q localisation in the cells of a silica granuloma. Granulomas were produced on the backs of adult albino rats by the subcutaneous injection of 1.0% suspension of Dornetrop quartz in saline. The granuloma illustrated was excised 12 d after its production; frozen sections were cut and C1q localised by the immunoperoxidase procedure[5]. There is a positive reaction around nuclei (clear holes) in many of the cells in the granulation tissue. It is possible that some of the reaction product is also extracellular, but it is difficult to resolve this point with certainty in this type of preparation.

In an attempt to demonstrate the origin of C1q *in vivo* various rat organs (including foetal and neonatal organs) were examined by the immunoperoxidase procedure as described above. This screen was entirely negative. C1q production, however, could be demonstrated *in vivo* in fibroblasts surrounding a silica granuloma which had been produced in the subcutaneous tissue of adult Wistar rats as shown in Fig. 3.

These data indicate that fibroblasts *in vitro* synthesise and secrete C1q. It seems likely that C1q is also produced by fibroblasts *in vivo* at least in pathological conditions where fibroblasts are functioning at a much higher rate than normal, namely in a silica granuloma. This evidence does does not exclude C1q production by other cell types *in vivo* or *in vitro*[7]. It is noteworthy, however, that freshly isolated human blood lymphocytes, the only cells we have found

which do not contain prolyl hydroxylase *in vitro* (and presumably do not synthesise collagen) do not contain C1q when examined by the immunoperoxidase procedure.

These results also suggest that primary fibroblast cultures may serve as a useful cell source for studies of C1q biosynthesis.

M. S. AL-ADNANI

The University Department of Pathology,
Royal Infirmary, Castle Street,
Glasgow, UK

J. O'D McGEE

Nuffield Department of Pathology, Radcliffe Infirmary,
University of Oxford, Oxford OX2 6HE, UK

1 Reid, K. B. M., Lowe, D. M., and Porter, R. R., *Biochem. J.*, **130**, 749 (1972).
2 Reid, K. B. M., *Biochem. J.*, **141**, 189 (1974).
3 Reid, K. B. M., and Porter, R. R., *Biochem. J.*, **155**, 19 (1976).
4 Nakane, P. K., and Pierce, G. P., *J. Cell Biol.*, **33**, 307 (1967).
5 Yonemasu, K., and Stroud, R. M., *J. Immun.*, **106**, 304 (1971).
6 Byers, P. H., McKenney, K. H., Lichtenstein, J. R., and Martin, G. R., *Biochemistry*, **13**, 5243 (1974).
7 Stecher, V. J., Morris, J. H., and Thorbecke, G. J., *Proc. Soc. exp. Biol. Med.*, **124**, 433 (1967).

Detection of α-foetoprotein in mouse liver differentiated hepatocytes before their progression through S phase

α-FOETOPROTEIN (AFP) is a serum protein present in large amounts during embryonic life and preserved in only nanogram quantities in adult mammals[1]. Increased levels of AFP are seen in adults in association with two forms of tumours —teratocarcinomas and hepatocarcinomas—and also during liver regeneration[2]. In early development AFP is synthesised in the yolk sac and in the liver[2,3] where it has been demonstrated by immunofluorescence in the majority of the hepatocytes[4]. Detailed aspects of AFP synthesis by hepatocytes, however, remain to be established. It has been suggested that it is synthesised by a definite population of hepatocytes[5], by one of the stages in their differentiation[6,7], or by liver cell precursors[8]. Sufficient data to decide between these alternatives are not presently available. We describe here the immunohistochemical localisation of AFP in regenerating mouse livers, in which DNA-synthesising cells have been identified by thymidine incorporation. AFP was found in a minority of adult differentiated hepatocytes, many of which had not progressed through S phase. Thus DNA synthesis is not a prerequisite for the induction of AFP synthesis in the adult liver.

Liver regeneration was induced by CCl_4 inhalation[9] in SWR or SWR/S mice (20–25 g). Serum AFP levels, as measured by immunodiffusion, increased significantly on day 2, reached a maximum on days 3–4 and had returned to basal levels by days 7–10. The centrolobular necrosis caused by CCl_4 resulted in practically all the remaining viable hepatocytes passing through one round of DNA synthesis, and a small fraction undergoing a second round. The most intensive DNA synthesis was found between 36 and 60 h and the duration of S phase in individual hepatocytes was 6.9 h (ref 10). These values were similar to those obtained in liver regenerating after partial hepatectomy, $S = 7.0 \pm 1.2$ h (ref. 11).

Mice with regenerating livers after exposure to CCl_4 were

Table 1 AFP in hepatocytes with and without thymidine incorporation

Mouse	1	2	3	4	5
% Labelled nuclei in total hepatocyte population	67 (^{3}H 55 + ^{14}C 12)	54 (^{3}H 42 + ^{14}C 12)	0.6 (^{3}H 0.3 + ^{14}C 0.3)	82	38
% Labelled nuclei in AFP-positive hepatocytes	78 (^{3}H 69 + ^{14}C 9)	34 (^{3}H 22 + ^{14}C 12)	4 (^{3}H 2.3 + ^{14}C 1.7)	82	64
Number of nuclei of AFP-positive hepatocytes examined	486	185	175	335	257

Mice exposed to CCl4[4] (ref. 11) were injected intraperitoneally with radioactive thymidine in 0.1 ml saline as follows. Mouse 1: 1 µCi g^{-1} methyl-^{3}H-thymidine (19.8 Ci mmol^{-1}, Isotope, USSR) every 6 h, with 1 µCi g^{-1} methyl-^{14}C-thymidine (39 mCi mmol^{-1}, Isotope, USSR) 1 h before being killed. Mice 2 and 3: 1 µCi g^{-1} methyl-^{3}H-thymidine (20.5 Ci mmol^{-1}, Isotope, USSR) every 3–4 h, with 0.8 µCi g^{-1} methyl-^{14}C-thymidine (54 mCi mmol^{-1}, Prague) 1 h before being killed. Mice 4 and 5: 1 µCi g^{-1} methyl-^{3}H-thymidine (20.5 Ci mmol^{-1}, Isotope, USSR) every 3-4 h, the last injection being given 1 h before being killed. Mouse 1 was killed 48 h after exposure and mice 2-5 after 42 h. Livers were fixed in cold ethanol–acetic acid[12]. AFP was localised with indirect immunofluorescence on 3-µm liver sections, using monospecific antibodies to murine AFP isolated from rabbit anti-murine AFP antiserum. Staining could be completely blocked by the addition of an equivalent amount of purified murine AFP (ref. 12). AFP-positive hepatocytes were photographed, after which the sections were washed, dehydrated, dried and covered with photosensitive emulsion (N ii Chim Photo, Moscow, USSR). After exposure for 1 month at 4 °C the autoradiographs were developed and the sections stained with haematoxylin–eosin. The percentage of nuclei labelled with ^{3}H or ^{14}C, distinguishable by grain density, were determined after counting 2,000 hepatocyte nuclei. Using the immunofluorescent photographs all AFP-positive cells were identified on corresponding autographs and nuclei with ^{3}H or ^{14}C label counted.

repeatedly injected with methyl-^{3}H-thymidine and killed by decapitation on day 2, 1 h after the last injection of radioisotope. In some mice the final injection was of methyl-^{14}C-thymidine, as described in Table 1. The livers were fixed immediately in cold ethanol–acetic acid and embedded in paraffin[12]. AFP was localised in liver sections by indirect immunofluorescence, and the nuclei of the cells in which DNA synthesis had occurred were subsequently identified by autoradiography. Damaged cells which could accumulate serum proteins nonspecifically because of increased membrane permeability were detected by the immunohistochemical demonstration of the serum marker protein, IgG. Consecutive sections were used so that individual cells could be stained for both AFP and IgG (ref. 12).

Using indirect immunofluorescence we failed to find any AFP-containing structures in the liver sections of the control normal adult mice. For the first time several AFP-positive hepatocytes were found in the liver section 24 h after the CCl$_4$ poisoning. During day 2 their amount and intensity of fluorescence increased rapidly, reached a peak on day 3 followed by a sharp decline on the day 4. On day 2, when the experiment was terminated, AFP was present in not more than 3–5% of the hepatocytes, in both mono- and binucleate cells. Most AFP-containing cells were found in the region adjacent to the centrolobular necrosis (Fig. 1), although in some mice they were present in the periportal area. Intense IgG fluorescence was observed in the necrotic centrolobular cells, demonstrating their uptake of serum proteins, but was absent in the great majority of AFP-positive cells. It does not therefore seem that the latter have stained because of nonspecific uptake of AFP due to elevated serum levels. In the following analysis of thymidine
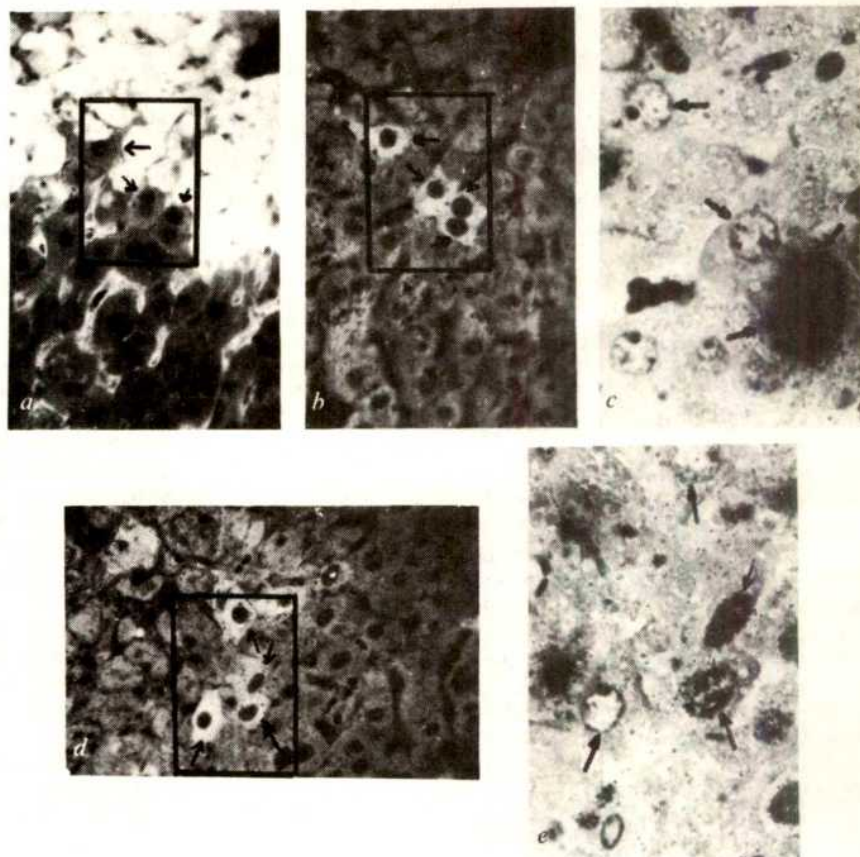
Fig. 1 AFP in the regenerating murine liver (mouse 1). Paraffin sections of liver fixed with ethanol–acetic acid and treated with rabbit anti-murine IgG serum (a) or purified rabbit antibodies anti-murine AFP (b and d) and FITC-conjugated donkey anti-murine IgG serum (Gamaleya Institute, Moscow, USSR). c and e, Autoradiographs of the areas ringed in b and d. Arrows show the same hepatocytes on the immunofluorescence photographs and on the autoradiographs. ^{14}C-label is easily distinguished from ^{3}H-label. b and c, Two nuclei of the AFP-positive hepatocytes are in S phase according to the intensive ^{14}C-label; two nuclei have no label. d and e, Two nuclei of the AFP-positive hepatocytes without label; two nuclei have incorporated ^{3}H-label and have progressed through S phase or all mitotic cycle. Magnification: a, b and d × 105; c and e × 225.

incorporation only IgG-negative, AFP-positive cells have been scored.

The percentage of labelled nuclei identified by autoradiography in the regenerating livers showed marked individual variation (Table 1). In all livers examined, however, numerous AFP-positive cells were seen without labelled nuclei (Fig. 1), as well as cells labelled in S phase with ^{14}C-thymidine 1 h before killing and cells labelled with ^{3}H-thymidine which had already progressed through S phase. The unlabelled cells comprised between 18 and 96% of the total of AFP-positive hepatocytes (Table 1).

The mice were injected with radioactive thymidine at intervals of 6 h (mouse 1) or 3–4 h (mice 2–5), that is less than the mean duration of S phase in these hepatocytes. As the thymidine was present in the circulation for at least 1 h after each injection, it is therefore unlikely that any DNA-synthesising cells could have escaped incorporation. Evidently AFP is localised in cells that have not yet commenced DNA synthesis during regeneration.

The time of appearance of the small number of AFP-positive hepatocytes, their quantity and intensity of fluorescence, all correlated directly with the changes in serum AFP measured. These observations, together with the failure of most AFP-positive cells to show nonspecific uptake of serum proteins, suggests that AFP is synthesised in these cells. If this is so, then the results show that AFP synthesis may be induced in the regenerating mouse liver in adult differentiated hepatocytes before the beginning of DNA synthesis. A similar suggestion has recently been made by Watanabe *et al.*[13], who have reported that serum AFP levels increase before the onset of proliferation in livers of rats treated with ethionine. In this context, it is worth noting that in an unrelated system (cultured erythroleukaemic cells), globin synthesis can be induced by butyric acid in the absence of cell division[14].

We thank Mrs M. D. Glishkina for assistance; Mrs V. S. Poltoranina, who started this work, Dr A. K. Yazova for monospecific antiserum to murine AFP and purified murine AFP; and the referee for comments.

N. V. ENGELHARDT

M. N. LAZAREVA

G. I. ABELEV

I. V. URYVAEVA

V. M. FACTOR

V. YA. BRODSKY

N.F. Gamaleya Institute for Microbiology and Epidemiology,
Institute of Developmental Biology,
Academy of Sciences,
Moscow, USSR

1 Ruoslahti, E., Pihko, H., and Seppala, M., *Transplant. Rev.*, 20, 38–60 (1974).
2 Abelev, G. I., *Adv. Cancer Res.*, 14, 295–358 (1971).
3 Gitlin, D., Boesman, M., *J. clin. Invest.*, 46, 1010–1016 (1967).
4 Gitlin, D., Kitzes, Y., and Boesman, M., *Nature*, 215, 534 (1967).
5 Abelev, G. I., *Cancer Res.*, 28, 1344–1350 (1968).
6 Onoé, T., Dempo, K., Kaneko, A., and Watabe, H., in *Alpha-fetoprotein and Hepatoma* (edit. by Hirai, H., and Miyaji, T.) (Gann Monograph, 14, 233–247, University Park Press, Baltimore, 1973).
7 Uriel, J., Aussel, C., Bonillon, D., de Nechaud, B., Loissillier, F., *Nature*, 244, 190–192 (1973).
8 Kitagawa, T., Yokochi, T., Sugano, H., *Int. J. Cancer*, 10, 368–381 (1972).
9 Bakirov, R. D., *Bull. exp. Biol. Med. (USSR)*, 2, 45–47 (1968).
10 Schultze, B., Gerhard, H., Shump, E., Maurer, V., *Virchow Arch. Abt.B., Cell Pat.*, 14, 329–343 (1973).
11 Factor, V. M., *Bull. exp. Biol. Med. (USSR)*, 9, 97–99 (1971).
12 Engelhardt, N. V., Goussev, A. I., Shipova, L.Ya., Abelev, G. I., *Int. J. Cancer*, 7, 198–206 (1971).
13 Watanabe, H., Miyazaki, M., Taketa, K., *Int. J. Cancer*, 17, 4, 518 (1976).
14 Leder, A., Opkin, S., Leder, P., *Science*, 190, 893–894 (1975).

Interaction of lighting and other environmental variables on activity of hypothalamo–hypophyseal–gonadal system

SINCE the pioneer study of Rowan[1], environmental lighting has been considered a predominant regulator of breeding activity in a variety of birds. A long daily photoperiod increases the levels of stored and circulating gonadotropins, and accelerates gonadal growth in Japanese quail[2,3], white-crowned sparrow[4], American tree sparrow[5], domestic pigeon[6] and mallard[7]. Other environmental variables such as temperature, rainfall and courtship are also known to influence breeding[8]. The nature of the endocrine response to temperature and rainfall is unknown, but the effect of male courtship on female endocrine systems has been studied extensively in ring doves (*Streptopelia risoria*). It increases the plasma levels of oestrogen[9], luteinising hormone[10] (LH) and progesterone[11], which in turn, induce sequential behavioural changes in the female that culminate in egg laying[12,13]. Using ring doves, I have investigated how environmental light and courtship interact to elicit the species-specific endocrine response. Since there is no significant change in testis size or weight during various stages of the breeding cycle of doves, this is not a suitable criterion for assessing the activity of the male endocrine state. Instead I used compensatory testicular hypertrophy, which is a response to environmental lighting in many species of birds, and investigated whether long daily photoperiod was required for this phenomenon in male ring doves, and whether courtship had any effect on this response.

Table 1 Description of experimental groups

Group	No. of males	Courtship condition	Daily light–dark (LD) cycle (h)
1	10	Isolation*	14:10
2	10	Isolation	8:16
3	10	Brief courtship†	14:10
4	10	Brief courtship	8:16
5	10	Breeding‡	14:10

* Males were housed individually in isolation rooms where they could hear, but not see, other birds.
† Males were introduced twice to a stimulus female in breeding cages that contained nest bowl and nest materials. Each exposure lasted 20 min. Sessions were spaced 1 week apart.
‡ Males were in a breeding cage that contained a stimulus female, nest bowl and nest materials throughout the 4-week experimental period.

Sixty reproductively experienced male ring doves from the indoor colony of the Institute of Animal Behavior, Rutgers University, were used. The colony is maintained under LD 14:10 controlled photoperiod and 21 ± 1 °C temperature. Fifty of the 60 birds were anaesthetised intramuscularly with equithesium (0.22 ml per 100 g), and the (larger) right testis of each was removed and weighed. The birds were then randomly assigned to one of five groups (Table 1), characterised in terms of daily photoperiod (LD 14:10 or LD 8:16) and the condition of courtship (isolation or breeding cage which contained a stimulus female). The remaining 10 doves were similarly anaesthetised, and both left and right testes removed and weighed. The stimulus females used for courtship condition were normal doves which showed sexual crouch, wing-flip and nest-coo[14].

Four weeks after dextral castration, all 50 males were laparotomised. The left testis (LT) of each was removed, and its weight expressed as a percentage of that of the right testis (RT). Since the LT weight of intact male doves kept under LD 14:10 was, on average, 80% that of the RT (Table 2), I considered LT $\geqslant 80\%$ of RT as evidence for compensatory

Table 2 Compensatory hypertophy of left testis after removal of right testis in the ring dove (*Streptopelia risoria*)

Group	Daily photostimulation/ courtship condition	No. of birds	Mean ± s.d. body weight (gm)	Mean ± s.d. right testis weight (mg)	Mean ± s.d. left testis weight (mg)	$\frac{\text{Left testis}}{\text{Right testis}} \times 100$ Mean	Range	Compensatory hypertrophy
1	14 h L Isolation	10	160±5	5,776±11	‡5,737±15	100†	85–114	Yes
2	8 h L Isolation	10	161±8	6,278±9	‡4,026±11	60*	30–74	No
3	14 h L Brief courtship	10	147±8	5,741±8	‡5,766±20	101†	96–109	Yes
4	8 h L Brief courtship	10	156±7	5,580±14	‡6,723±26	112†	94–130	Yes
5	14 h L Breeding	10	159±6	5,780±9	‡5,323±32	96	83–120	Yes
C	14 h L Isolation control	10	158±6	5,661±12	4,632±17	80	75–86	

Analysis of variance for the randomised groups: $F = 3.52$, $\alpha = 0.05$.
* *t* test significant ($P < 0.01$) only between group 2 and every other group.
† *t* test significant ($P < 0.01$) in comparison with group C.
‡ Measurements were taken 4 weeks after dextral castration.

hypertrophy. The results are summarised in Table 2. (1) Compensatory testicular hypertrophy occurred in birds under LD 14:10 (groups 1, 3 and 5), but not in birds kept isolated and under LD 8:16 (group 2); (2) LT weights of birds in group 2 averaged only 60% that of RT, indicating that the LT had regressed rather than augmented. (An histological study showed an increase in size of seminiferous tubules, Leydig cells and sperm count in the augmented testis.)

These results, together with the finding that LD 8:16 caused testicular regression in ring doves (unpublished) and the fact that other species of birds show a decline in plasma LH level[3-5], demonstrate that long daily photoperiod is an important regulating factor for the feedback function of the hypothalamo–hypophyseal–gonadal system. Short exposure to courtship (two pairing sessions with stimulus females for 20 min each and 1 week apart) however, resulted in good testicular hypertrophy in male doves kept under LD 8:16 (group 4). The stimulatory effect of courtship on male doves does not seem to increase with increased duration of exposure to stimulus females. There was no apparent difference in LT weight among birds kept under LD 14:10, regardless of whether the birds were isolated or exposed to stimulus females (groups 1, 3 and 5). Thus, it seems that the combined effects of courtship and lighting on the hypothalamo–hypophyseal–gonadal system are guided by an adaptive change of endocrine state for the species, rather than simply a mathematical sum of respective environmental effects.

Two brief exposures to a stimulus female (20 min each and 1 week apart, group 4) affected males in a manner comparable with that of 6 h additional daily photoperiod for 4 weeks (group 1); this makes the stimulatory effect of female courtship on compensatory testicular hypertrophy even more significant. This is equally true of the effect of male courtship on female systems.

Under long photoperiod (LD 14:10), follicles of female ring doves took 20 weeks to develop from 2–3 mm to 8–9 mm in diameter, and they then became atresic. In contrast, when paired with male doves in breeding cages, female ring doves of similar ovarian stage (follicle size 2–3 mm) took only 7–11 d to develop large follicles (14–15 mm) which subsequently ovulated and oviposited (unpublished).

In male and female Japanese quail, although long photoperiod increases gonadotropin-releasing factor in the posterior basal hypothalamus area[3,15], there is no direct evidence to suggest a similar effect of courtship stimulation on releasing factor by the hypothalamus. In my study, compensatory hypertrophy in birds kept under LD 8:16 and exposed to stimulus females twice (group 4), strongly suggests that courtship stimulates either: gonadotropin synthesis, gonadotropin release, or the sensitivity of the pituitary to gonadotropin-releasing factors. I favour the second possibility in view of the following findings in the

Japanese quail: (1) Long daily photostimulation results in gonadal development[3] and hypothalamic gonadotropin-releasing activity[15]; and (2) a change of daily photoperiod from LD 14:10 to LD 8:16 results in a lower level of plasma LH and testicular regression, but a higher level of pituitary LH; these events occurred within 6–16 d following the onset of short photoperiod[3]. Similarly, in male golden hamsters, exposure to non-stimulatory photoperiod does not alter the increased pituitary LH and follicle stimulating hormone (FSH) content after castration, whereas the serum levels of LH and FSH decline substantially[16].

On the basis of these findings, it seems that environmental photoperiod affects the synthesis as well as release of gonadotropins; the releasing mechanism being more responsive to an immediate photoperiodic change than the synthesis, whereas synthesis (and/or store) is regulated by a long term photoperiodic change and endogenous feedback system. The failure of the hypertrophy response in birds kept in short light regimes (group 2) is thus a result of the cessation of gonadotropin release in response to the short photoperiod. Gonadotropin release can, however, be induced by other environmental stimulation (such as a female stimulus, warm temperature, rainfall and so on), which results in the hypertrophy seen in group 4 birds kept in short light regimes. The release is possible since synthesis (and/or store) of gonadotropin is not affected by immediate photoperiodic change.

Mei-Fang Cheng

Institute of Animal Behavior,
Rutgers University,
101 Warren Street,
Newark, New Jersey 07102

1 Rowan, W., *Proc. natn. Acad. Sci. U.S.A.*, 16, 520–525 (1930).
2 Follett, B. K., and Farner, D. S., *Gen. comp. Endocr.*, 7, 111–124 (1966).
3 Nicholls, T. J., Scanes, C. G., and Follett, B. K., *Gen. comp. Endocr.*, 21, 84–98 (1973).
4 Follett, B. K., Farner, D. S., and Mattocks, P. W., Jr, *Gen. comp. Endocr.*, 26, 126–134 (1975).
5 Wilson, F. E., and Follett, B. K., *Gen. comp. Endocr.*, 23, 82–93 (1974).
6 Bhattacharyya, T. K., and Mujushri, S., *Acta morph. A. hung.*, 17, 113–122 (1969).
7 Benoit, J., *N.Y. Acad. Sci.*, 117, 23–34 (1964).
8 Marshall, A. J., in *Biol. and Comparative Physiology of Birds* (edit. by Marshall, A. J.), 307–332 (Academic, New York and London, 1961).
9 Korenbrot, C. S., Schomberg, D. W., and Erickson, C. J., *Endocrinology*, 94, 1126–1134 (1974).
10 Cheng, M. F., and Follett, B. K., *Hormone. Behav.*, 7, 1–7 (1976).
11 Silver, R., Reboulleau, C., Lehrman, D. S., and Feder, H. H., *Endocrinology*, 94, 1547–1554 (1974).
12 Cheng, M. F., *J. comp. Physiol. Psychol.*, 83, 234–239 (1973).
13 Cheng, M. F., and Silver, R., *J. comp. Physiol. Psychol.*, 88, 256–263 (1975).
14 Cheng, M. F., *J. comp. Physiol. Psychol.*, 83, 221–233 (1973).
15 Follett, B. K., *Gen. comp. Endocr.*, 15, 165–179 (1970).
16 Turek, F. W., Elliot, J. A., Alvis, J. D., and Menaker, M., *Endocrinology*, 96, 854–860 (1975).

A discontinuous relationship between the acetylcholine-activated channel conductance and temperature

SINCE the introduction of acetylcholine (ACh) 'noise' analysis[1,2], it has been shown that several manipulations affect ACh-receptor molecular events[3-5]. The ACh receptor is an integral membrane protein and is therefore intimately associated with surrounding membrane lipids. We have begun experiments using embryonic chick muscle in culture to determine whether surrounding lipids affect the function of the ACh receptor. As a first step, we have analysed ACh current noise to determine the temperature dependence of the mean single channel conductance ($\bar{g}_o$) and open time ($\bar{t}_o$).

At frog endplates $\bar{t}_o$ is gradually prolonged as the temperature (T) is lowered but $\bar{g}_o$ remains essentially unchanged. We observed a similar relation between $\bar{t}_o$ and T in cultured chick muscle, but in contrast to results obtained in amphibia, $\bar{g}_o$ was markedly decreased at low temperatures. Moreover, there was a marked discontinuity in the relation between T and $\bar{g}_o$. A clear 'transition' temperature was observed at ~ 20 °C which suggests that $\bar{g}_o$ may depend on the 'fluidity' of the receptor microenvironment.

ACh-induced transmembrane currents were recorded in voltage clamped large spherical muscle cells (Fig. 1). Such spheres or 'myoballs' are more suitable than elongated cylindrical muscle fibres for voltage clamp analysis. The interior of each myoball is isopotential and the membrane potential can be held constant, even in the face of large ACh responses (up to 200 nA). Myoballs are 'normal' in the sense that they are extremely sensitive to ACh and they contract.

Control of membrane potential was achieved with two intracellular electrodes—one for measuring the membrane potential and the other for supplying the feedback current required to hold the membrane potential constant. The microelectrodes had a resistance of 3–10 MΩ when filled with 3 M KCl. A microelectrode filled with ACh was located 50–100 μm away from the cell and ACh was ejected by prolonged, positive pulses. Temperature was controlled by perfusing ethanol (+ 60 °C to − 30 °C) between the glass bottom of the recording chamber and the objective of the inverted microscope. In most

experiments, the membrane potential was clamped at − 80 mV. Feedback currents were stored on magnetic tape (frequency response 0–1,250 Hz).

A typical ACh response is shown in Fig. 2a. The horizontal bar indicates the duration of the ACh pulse. The lower trace is a low gain d.c. recording of the membrane ionic current and the upper trace is a parallel, high gain filtered (bandpass 1–200 Hz) recording which shows the fluctuations or noise of the ACh-induced current. In this example, the maximum inward current produced by ACh was 100 nA.

ACh responses were analysed with a PDP/8e computer. The single-channel conductance, $\bar{g}_o$, was estimated from the slope

a

b
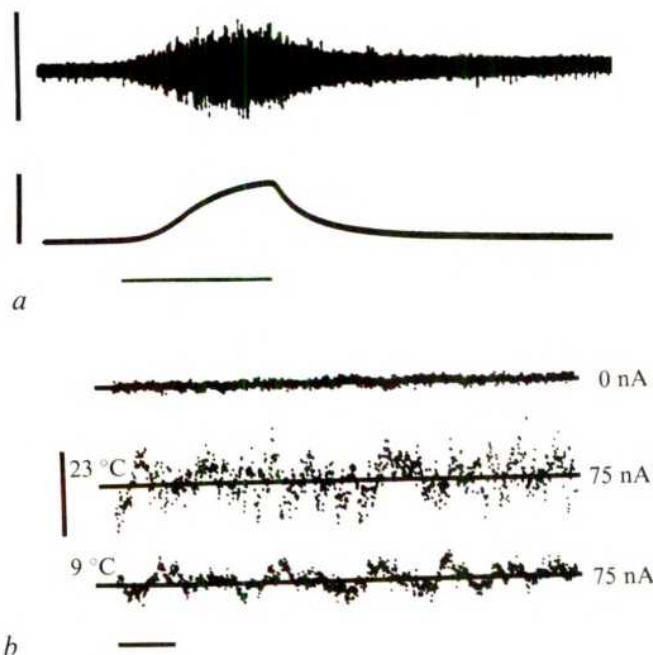
Fig. 2 ACh-induced membrane currents in voltage clamped myoballs. *a*, A 15-s long pulse of ACh (horizontal bar) results in a large increase in mean membrane current (lower trace: a low-gain d.c. recording) and an increase in current fluctuations (upper trace: a parallel high gain filtered channel). The membrane potential was clamped to −95 mV. The slow rise of the response reflects the time required for ACh to diffuse from the ACh pipette located 50 μm away. Vertical bars: 3 nA upper trace, 100 nA lower trace. *b*, Digitalised, 2-s long, samples of ACh current fluctuations (noise). The upper trace shows control fluctuations in the absence of ACh. The steady holding current is defined as zero. ACh responses obtained at 23 and 9 °C are shown in the middle and lower traces respectively. The mean ACh current was 75 nA at both temperatures. Note that at 9 °C, high frequency components are less prominent and that the amplitude of the noise is decreased. Vertical bar, 2 nA; horizontal bar = 250 ms.

Fig. 1 A large chick myoball. Spheres were obtained by adding mononucleated muscle precursor cells dissociated from 11-d pectoral muscles to Petri dishes coated with Sylgard (Dow Corning). The cells did not attach to this surface, and over the next 2–3 d muscle precursor cells fused with one another in suspension to form multinucleated 'myoballs', 50–120 μm in diameter. For electrophysiological experiments, myoballs were allowed to settle on collagen-coated glass coverslips which were then placed in a chamber on an inverted phase contrast microscope. Smooth surfaced myoballs could be easily distinguished from aggregates of fibroblasts which had a cobblestone-like appearance. Bar = 100 μm.

of the linear relation between the variance of current fluctuations and the mean membrane current. The mean channel open time, $\bar{t}_o$, was estimated from the spectral density function, $S(f)$; $\bar{t}_o = 1/2\pi f_c$ where f_c is that frequency at which $S(f)$ is reduced by half[1,2,6]. At 37 °C, when the membrane potential was set between −70 and −95 mV, $\bar{g}_o$ ranged between 25 and 45 pmho, and $\bar{t}_o$ between 2.5 and 4.0 ms.

Figure 2b shows digitalised records of control or background noise and ACh-induced noise at two temperatures (sampling rate: 500 s⁻¹). As reported in previous studies in amphibia, lower frequency components predominate at the lower temperature. In this example, $\bar{t}_o$ increased from 7.2 ms at 23 °C to 16.1 ms at 9 °C. In these chick cells, $\bar{t}_o$ increased continuously as T was lowered from 37 to 7 °C. In addition, Fig. 2 shows that the amplitude of the current fluctuations is markedly decreased at 9 °C, even though the mean ACh-induced current is the same at the two temperatures. Estimates of $\bar{g}_o$ in this cell decreased
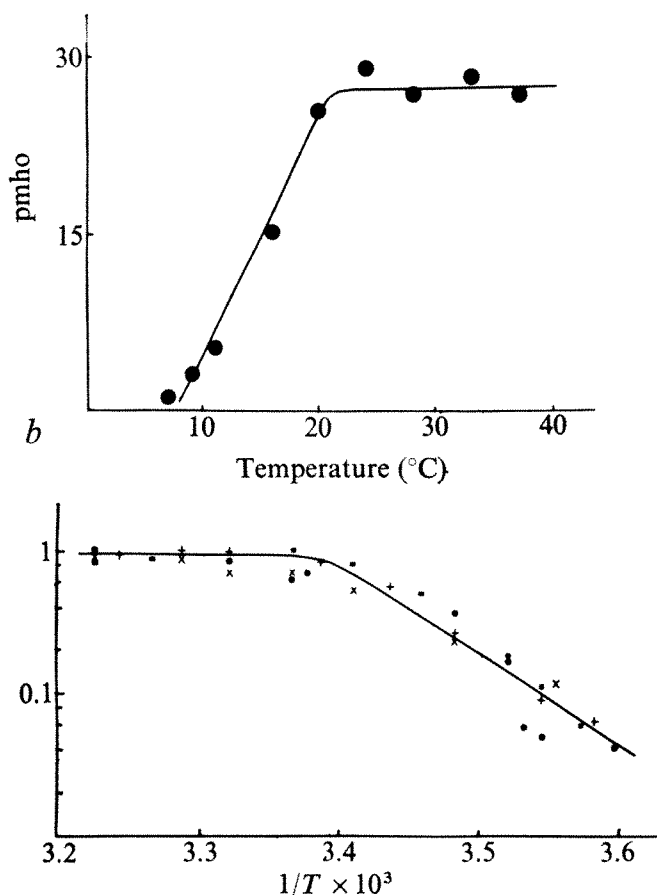
Fig. 3 Temperature dependence of single channel conductance $\bar{g}_o$ in chick myoballs. *a*, Nine ACh responses elicited in the same cell at different temperatures between 7 and 37 °C are shown. The relationship between $\bar{g}_o$ and temperature is discontinuous; a sharp transition is observed at ~ 20 °C. *b*, Arrhenius plot of normalised data obtained in four different myoballs. The activation energy (E_a), calculated from the equation $\log g_o = C - (E_a/2.3R)(1/T)$ (R = gas constant, T = absolute temperature, and C = a constant) increases from virtually 0 above 20 °C to 30 kcalorie mol^{-1} below this critical temperature.

in this report. The degree to which an ACh-activated channel opens may depend on the flexibility of fatty acids in its micro-environment. It may be significant that the membranes of poikilothermic animals (in which $\bar{g}_o$ remains nearly constant between 10 and 23 °C) contain relatively more unsaturated fatty acids which 'melt' at a lower temperature[9]. The fact that no clear break was observed in the relation between $\bar{t}_o$ and T may simply mean that different molecular processes regulate $\bar{t}_o$ and $\bar{g}_o$.

It should be possible systematically to modify the membrane lipids of rapidly growing muscle cells by adding specific fatty acids to a culture medium made up with delipidated serum[10]. The relationship between $\bar{g}_o$ and membrane fluidity might then be defined in more detail. In addition, changes in regional and/or overall membrane fluidity may alter the degradation of ACh receptors or their insertion into, or distribution within, the membrane.

We thank C. F. Stevens for help with the voltage clamp amplifier and S. M. Schuetze for the computer programming. This investigation was supported in part by a Fogarty International Research Fellowship and Yad Avi Ha-Yishuv Foundation, Israel (Y.L.), and by a USPHS Research Grant from the NIH (G.D.F.).

Yoram Lass*
Gerald D. Fischbach

Department of Pharmacology,
Harvard Medical School,
Boston, Massachusetts 02115

1 Katz, B., and Miledi, R., *J. Physiol., Lond.*, **224**, 665–699 (1972).
2 Anderson, C. R., and Stevens, C. F., *J. Physiol., Lond.*, **235**, 665–691 (1973).
3 Landau, E. M., and Ben-Haim, D., *Science*, **185**, 944–946 (1974).
4 Katz, B., and Miledi, R., *J. Physiol., Lond.*, **249**, 269–284 (1975).
5 Gage, P. W., *Physiol. Rev.*, **56**, 177–247 (1976).
6 Bendat, J. S., and Piersol, A., *Random Data: Analysis and Measurement Procedures* (Wiley, New York, 1971).
7 Overath, P., and Trauble, H., *Biochemistry*, **12**, 2625–2634 (1973).
8 Linden, C. D., Wright, K. L., McConnell, H. M., and Fox, C. F., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 2271–2275 (1973).
9 Lehninger, A. L., *Biochemistry*, 190 (Worth, New York, 1970).
10 Horwitz, A. F., Hatten, M. E., and Burger, M. M., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 3115–3119 (1974).

from 31 pmho at 23 °C to 7.6 pmho at 9 °C. A comparable or larger decrease in $\bar{g}_o$ was obtained whenever the temperature was lowered to 10 °C, and in each case the change in $\bar{g}_0$ was completely reversed on rewarming the preparation. This dramatic effect cannot be explained by a negative shift in the ACh reversal potential (measured by ACh application at membrane potentials between -80 and $+30$ mV) which remained fixed at ~ -5 mV between 37 and 5 °C.

In contrast to the continuous relationship between $\bar{t}_o$ and T, there was a clear discontinuity in the relation between $\bar{g}_o$ and T in every case examined (Fig. 3). There was no significant decrease in $\bar{g}_o$ as T was lowered from 37 to 20 °C. As T was lowered further to 5 °C, however, there was a nearly tenfold decrease in $\bar{g}_o$. An Arrhenius plot including points from four myoballs is shown in Fig. 3*b*. From the slope it can be determined that the activation energy for those factors or events which determine the magnitude to which an ACh channel can open is increased by 30 kcalorie mol^{-1} at ~ 20 °C.

A transition temperature has been observed in studies of other membrane proteins where enzyme activity, or transport or mobility within the lipid bilayer were assayed[7,8]. It is thought that a change in membrane fluidity occurs at the critical temperature. That is, above the critical temperature, the phospholipid fatty acid side chains are flexible and the interior of the membrane is fluid; below the critical temperature, movements of fatty acid side chains are restricted and the membrane core is more solid. It is plausible that the same mechanism explains the discontinuous relation between $\bar{g}_o$ and T described

Long term potentiation is accompanied by a reduction in dendritic responsiveness to glutamic acid

Hippocampal synapses show an unusual degree of physiological plasticity. Relatively modest levels of repetitive stimulation cause considerable enhancement in subsequent responses to single pulse stimulation, and this potentiation lasts for hours and even days[1–6]. These findings, which have now been replicated in several laboratories, have aroused much interest, first because they seem to represent an excellent starting point for the analysis of synaptic plasticity in mammalian brain, and second, because their rapid development and persistence suggest that they may be related to processes involved in behavioural plasticity. The studies reported here represent an attempt to analyse the mechanisms underlying physiological plasticity in the hippocampus, in particular the locus and anatomical specificity of the effect. The change in magnitude of response which occurs after repetitive stimulation could reflect either presynaptic or postsynaptic adjustment. If the latter were the case, the change might involve a particular dendritic region innervated by the stimulated input, or alteration in the status of the entire postsynaptic neurone. Seeking data relevant to these questions, we have measured the effects of glutamic acid applied iontophoretically to different levels
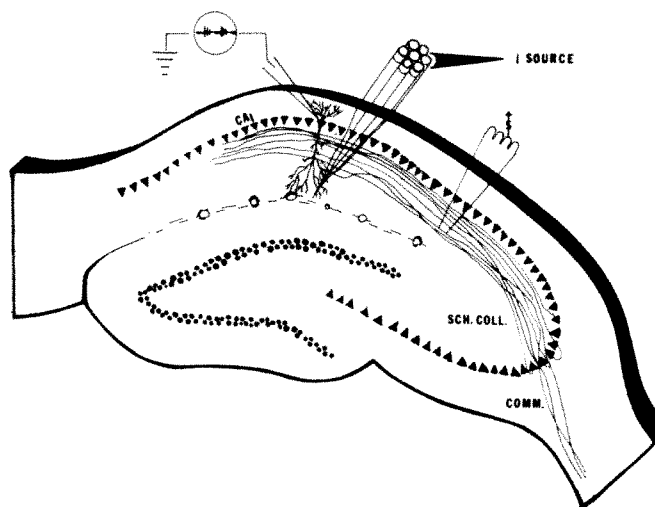
Fig. 1 Schematic representation of the transverse hippocampal slice and the experimental electrode arrangement. Extracellular unit activity and evoked potentials were recorded from the CA 1 pyramidal cell somal layer as indicated. The separate multi-barrelled iontophoresis electrode could be placed at any point in the apical and basal dendritic zones of a particular neurone for the purpose of glutamic acid ejections. Schaffer collateral–commissural stimulation was accomplished through a bipolar electrode placed along the tract at the CA 1–CA 2 border.

of the dendritic trees of the pyramidal cells before and after potentiation of the extracellular postsynaptic response.

Glutamate more or less universally excites neurones[7-9] and it is generally thought that it does this by depolarising the cell through an action on sodium and possibly potassium channels[9]. We reasoned that if potentiation produced a lasting change in the membrane properties of the post-synaptic cell, this might cause changes in that cell's response to discrete applications of glutamic acid. Furthermore, if the postsynaptic effect were generalised, that is, not restricted to the dendritic area affected by the driven input, then an altered response would be found to glutamic acid applied anywhere in the dendritic field. To test these possibilities we have used *in vitro* transverse slices of rat hippocampus, prepared and maintained as described before[10]. The experimental arrangement is described in Fig. 1. Briefly, a bipolar stimulating electrode (62 μm wire) was placed on the Schaffer collaterals (the fibre connection consisting of collateral axons from the CA 3 field to the CA 1 field of pyramid cells (Fig. 1)) and a recording micropipette was guided into the layer of pyramidal cell bodies, the proximal dendrites of which receive the Schaffer input. The recording electrode was advanced until it was located next to a spontaneously active cell. The electrode recorded both the spike activity of the "isolated" cell and the characteristic field potential elicited by stimulation of the Schaffer collaterals. This response is a simple monophasic positive wave when threshold stimulation (4–7 V) is used; at slightly higher voltages a sharp negative deflection is superimposed on the wave (Fig. 2). This negative wave has been shown by Andersen and associates to be an envelope of synchronously driven cell spikes, that is, in their words, a "population spike"[11]. We used this response as our index of the magnitude of the postsynaptic response. A third multibarrelled electrode was used for iontophoretic application of 0.2 M L-glutamate to various regions of the pyramidal cells' dendritic field. Only a single CA 1 dendritic locus was tested in each slice. When the recording and electrical stimulation electrodes were in position the iontophoresis electrode was lowered into the chosen region of the dendritic field while glutamate was ejected at 20–30 nA until the recorded cell was vigorously, though not maximally, driven. The iontophoresis current was then lowered until a short latency and stable response were obtained for a 10-s test period. When all three electrodes were in position the experiment was

begun (usually at about 2 h after preparation of the slices). A constant current source was programmed to pulse the glutamate on to the CA 1 dendrites for 10 s every 40 s for 3–5 min. During this time averages of 10 field potentials were collected in response to stimulation of the Schaffer collateral bundle. Stimulus pulses were applied once every 5 s at a voltage adequate to elicit a small (less than 1-mV) population spike. Then 15-Hz tetanic stimulation was delivered for 15 s, after which iontophoretic applications were conducted as before. Field potentials were again collected and averaged 1, 4, 7, 12, and 20 min after 15-Hz stimulation —these were elicited by the same voltage and frequency as used before the conditioning train.

Twenty-two slices were tested by this procedure. Four of these did not show potentiation of the "population spike"
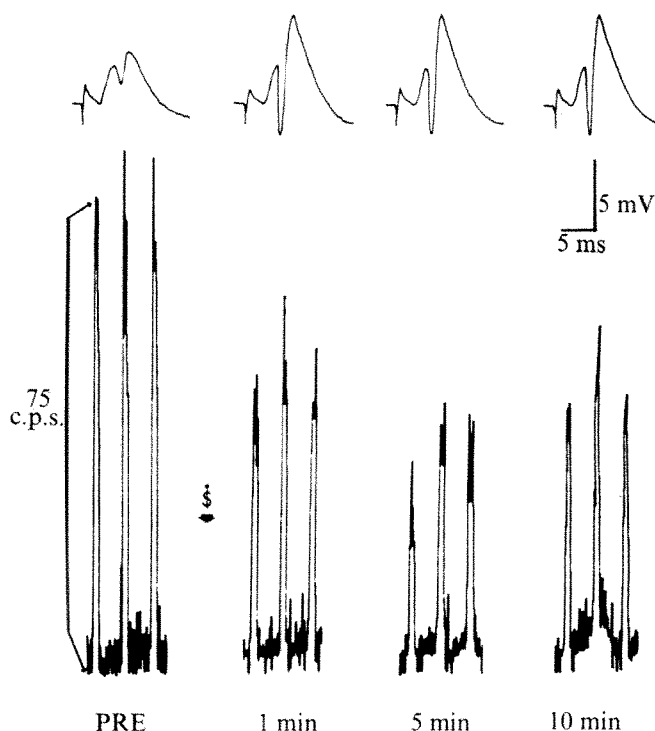
Fig. 2 The upper traces are computer averages of the field potential recorded by the pipette located in the cell body layer before (Pre) and 1, 5 and 10 min after a 15-Hz, 15-s train of stimulation pulses (denoted by S) to the Schaffer collateral system. The field potential and the population spike, the negative deflection riding on the positive wave, were greatly enhanced by the tetanising stimulation. The lower traces are records of the rate of spike discharges of a small population of pyramid cells, recorded by the same microelectrode used to collect the field potentials, in response to a 10-s iontophoretic pulse (25nA) of glutamic acid applied in apical dendritic field 200 μm away from the cell layer. Repetitive stimulation caused a decrease in the excitation produced by the amino acid.

after the conditioning pulses and in two others the degree of response facilitation was quite modest. In the rest of the slices, the "post-tetani" responses were at least twice the amplitude of the control responses. In a few slices from this last group the potentiation was transient and the response was clearly returning towards baseline by the end of the test runs.

The effects of the tetanising stimulation on the unitary response to glutamic acid were consistently the reverse of that displayed by the field potential. In addition it seemed that the nature and degree of the decrease in response to glutamic acid were related to the duration and magnitude of field potential potentiation. In slices which showed a potentiated response the excitation elicited by glutamic acid was reduced by at least 25%. The effect appeared within 30 s of the termination of the conditioning train and

generally persisted throughout testing. Furthermore, in those slices in which the potentiation was transient the reduction of the response to glutamic acid gradually dissipated (Fig. 3).

As mentioned, two slices showed very modest potentiation and in these cases the suppression of glutamic acid excitation was less than 25%. Finally, three of four slices which did not potentiate also did not exhibit any change in responsiveness to iontophoretically applied glutamic acid. The one exception demonstrated a very slight and transient reduction. Table 1 summarises these points. In 10 of the slices we used a second train of conditioning pulses and measured the effects of this on the response of the pyramidal cells to glutamic acid. The results of these tests were comparable with those obtained after the first tetanus: that is a strong potentiation ($N=6$) was accompanied by a reduction in responsiveness to glutamic acid, whereas absence of potentiation ($N=4$) was correlated with essentially no change in iontophoretically induced unitary excitation (Table 1, bottom).

Glutamate has been found (ref. 12 and unpublished results of V.K.G., H. J. Spencer, C. W. Cotman and G.S.L.) to excite pyramidal cells more or less independently of where it is applied to the dendritic tree, and our study confirmed this. More pertinent to our findings, depression of glutamic acid-induced excitation was found even when the amino acid was applied outside the region of Schaffer collateral innervation (to the basal dendrites), or to the apical dendritic tips in the zone of entorhinal innervation (Fig. 1).

Fig. 3 An example of post-tetanic potentiation which did not persist for more than 15 min after the 15-Hz, 15-s conditioning train (S). Note the reduction of glutamic acid effect while the field response was in a potentiated state, and the return of driving to preconditioning levels as the evoked potential decreased to baseline levels. All variables are the same as in Fig. 2.
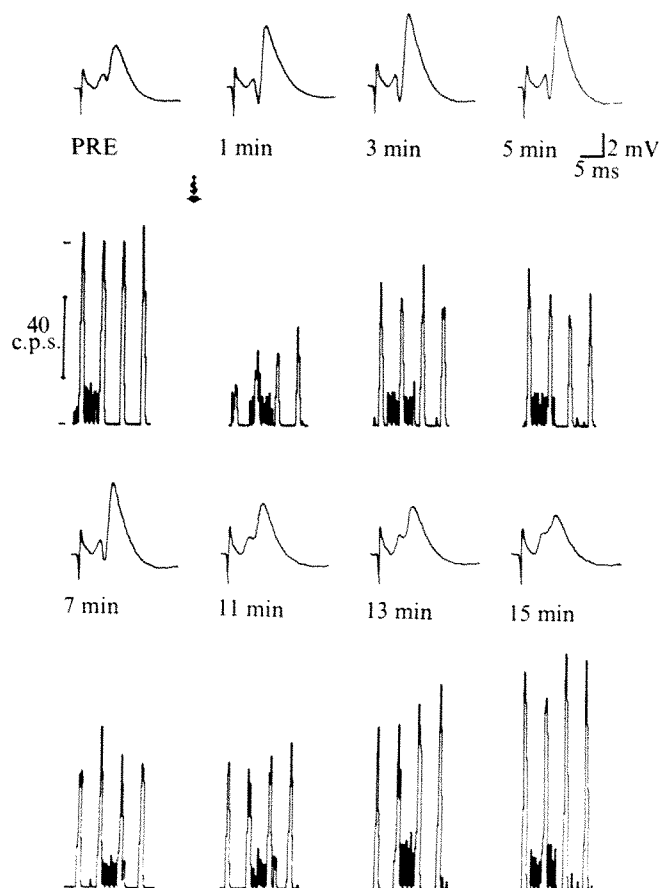
Table 1 Potentiation and response to glutamate

		Relative degree of potentiation			
		0	1	2	
Relative depression of responses to glutamate	0	3	—	—	
	1	1	2	1	First
	2	—	—	8	test
	3	—	—	5	
	0	2	—	—	
	1	1	1	—	Second
	2	—	—	3	test
	3	—	—	3	

The figures are the numbers of slices displaying the noted degree of potentiation (0, no change; 1, less than 100% increase in population spike within 1 min of the end of stimulation train; 2, greater than 100% increase in population spike in this time frame) and a particular degree of reduction in response to iontophoretic application of glutamic acid (0, no change; 1, 10–25% reduction of firing rate; 2, 25–40% reduction in firing rate; 3, greater than 40% reduction). First and second tests are explained in text.

These results indicate that tetanic stimulation produces a lasting change in responsivity of the postsynaptic membrane and that this change is generalised to the entire neurone. The similarity in degree and time course of the alteration in glutamic acid responsiveness and the size and duration of potentiation suggest that these two phenomena are related. Postsynaptic depression after tetani has been reported in the peripheral nervous system (for example ref. 13) but the rapid development and persistent nature of the effects reported here make it unlikely that comparable processes are involved. Speculation as to the physiological nature of the changes responsible for the altered glutamic acid response will be more profitable after completion of intracellular recording studies which are currently in progress.

G. S. Lynch
V. K. Gribkoff
S. A. Deadwyler

*Department of Psychobiology,
University of California,
Irvine, California 92717*

1 Lomo, T., *Expl Brain Res.*, **12**, 46–63 (1971).
2 Bliss, T. V. P., and Lomo, T., *J. Physiol., Lond.*, **232**, 331–356 (1973).
3 Bliss, T. V. P., and Gardner-Medwin, A. R., *J. Physiol., Lond.*, **232**, 357–374 (1973).
4 Douglas, R. M., and Goddard, G. V., *Brain Res.*, **86**, 205–215 (1974).
5 Schwartzkroin, P. A., and Wester, K., *Brain Res.*, **89**, 107–119 (1975).
6 Deadwyler, S. A., Gribkoff, V., Cotman, C. W., and Lynch, G., *Brain Res. Bull.*, **1**, 1–7 (1976).
7 Curtis, D. R., Phyllis, J. W., and Watkins, J. C., *J. Physiol., Lond.*, **150**, 656–682 (1960).
8 Krnjevic, K., and Phyllis, J. W., *J. Physiol., Lond.*, **165**, 274–327 (1963).
9 Krnjevic, K., *Physiol. Rev.*, **54**, 418–540 (1974).
10 Lynch, G., Smith, R. L., Browning, M. D., and Deadwyler, S., in *Adv. Neurol.*, **12**, (edit. by Kreutzberg, G.), 297–313 (Raven, New York, 1975).
11 Andersen, P., Bliss, T. V. P., and Skrede, K. K., *Expl Brain Res.*, **13**, 208–238 (1971).
12 Spencer, H. J., Gribkoff, V. K., Cotman, C. W., and Lynch, G., *Brain Res.*, **105**, 471–481 (1976).
13 Thesleff, S., *J. Physiol., Lond.*, **148**, 659–664 (1959).

Inhibitory postsynaptic current in voltage-clamped crayfish muscle

THE current–voltage relationship provides important information about the mechanism of ionic permeation through membranes. In excitatory synapses the relationship between membrane potential and excitatory postsynaptic current (e.p.s.c.) has been extensively studied using the voltage-clamp method[1-6]. In the frog endplate it has been shown that the e.p.s.c. was smaller than would be expected from a linear

relationship during hyperpolarisation of the membrane[3,4]. On the other hand, after iontophoretic application of acetylcholine, the synaptic current increased more than expected at hyperpolarised membrane potentials[7]. In investigating the current–voltage relationship it is important to measure the current on both sides of the equilibrium potential. Inhibitory postsynaptic current (i.p.s.c.) is more suitable for this purpose, because the reversal potential is near the resting potential and so the current–voltage relationship on both sides of the reversal potential can be measured using relatively small potential changes. Little attention has been paid, however, to inhibitory postsynaptic membranes, possibly because the inhibitory synapses are distributed diffusely over the surface of the cell or situated on dendrites remote from the cell body where the electrical changes are measured. In these conditions it is difficult to measure accurately the current–voltage relationship. We have cannulated the opener muscle of the claw in the crayfish (*Cambarus clarkii*) using a stainless steel wire, the membrane potential being clamped at various levels as described previously[6]. Although the inhibitory synapses are distributed over the whole surface of the muscle fibre, the space clamp condition of the muscle fibre was maintained well using this technique. We have found that the relationship between membrane potential and i.p.s.c. was highly nonlinear on hyperpolarisation of the membrane.

Fig. 2 Relationship between clamped membrane potential and amplitude of i.p.s.c. *a*, Peak amplitude of i.p.s.c. produced at 16 Hz; *b*, plateau level of i.p.s.c.s produced at 107 Hz. Inside current expressed as positive.
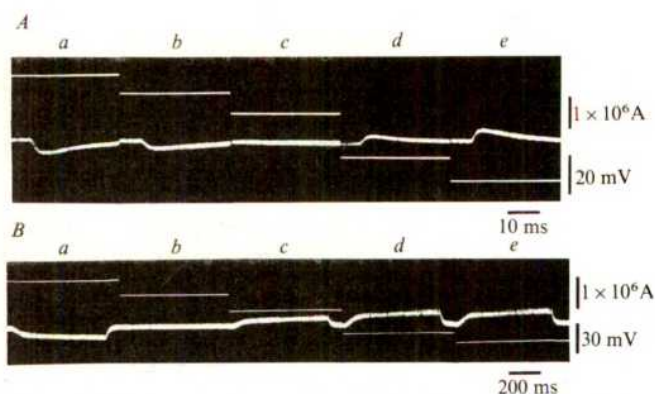
Fig. 1 Inhibitory postsynaptic current (i.p.s.c.) recorded at various membrane potentials. *A*, Superimposed i.p.s.c. at stimulation rate of 16 Hz: *a*, −37 mV; *b*, −46 mV; *c*, −57 mV; *d*, −79 mV; *e*, −90 mV. Upward deflection indicates the inside current. *B*, Train of i.p.s.c.s produced at 107 Hz: *a*, −47 mV; *b*, −60 mV; *c*, −76 mV; *d*, −99 mV; *e*, −109 mV.

The inhibitory axon was dissected at the melopodite and stimulated using a pair of silver electrodes. Inhibitory stimulation produced the inwardly directed i.p.s.c. at hyperpolarised membrane potentials (Fig. 1*A*, *d* and *e*). Although the time course of the i.p.s.c. was rather variable from fibre to fibre, the current reached its peak within ∼ 4–5 ms and thereafter decayed almost exponentially with a time constant of 10–20 ms (total duration up to 50 ms). On depolarisation of the membrane, the amplitude of the i.p.s.c. decreased and the current reversed its sign at about −60 mV (Fig. 1*A*, *c*). The peak time and falling phase of i.p.s.c. at depolarised membrane potentials (downward deflections in Fig. 1*A*, *a* and *b*) were slightly longer than those at hyperpolarised membrane potentials. The time constant of the falling phase was 15.6 ± 5.1 ms at about −100 mV and 21.3 ± 6.0 ms at −40 mV (mean ± s.d., $n = 8$).

The time course of changes in i.p.s.c. was much slower than that of the e.p.s.c. recorded from the same muscle; the peak time of e.p.s.c. being about 2–3 ms and the time constant of the falling phase being 3.3 ± 1.1 ms at −100 mV. On repeated stimulation of the inhibitory axon, the falling phases of the i.p.s.c. summed up to a plateau (Fig. 1*B*). Such a summation was, however, not manifest in the e.p.s.c. (ref. 6.)

The peak amplitude of a single i.p.s.c. (Fig. 2*a*) and the plateau level of repeated i.p.s.c.s (Fig. 2*b*) were plotted against the clamped membrane potential. The synaptic current–membrane potential relationship was almost linear for the depolarised membrane up to −30 or −15 mV. When the membrane potential was hyperpolarised by about 20–30 mV from the reversal potential, however, the relationship deviated from linearity and the i.p.s.c. reached an approximately limiting value (Fig. 2*b*). This is different from observations of e.p.s.c., which increased almost linearly up to −120 to −150 mV (ref. 6). Since the i.p.s.c. is carried by Cl⁻ (ref. 8), the outwardly directed Cl⁻ flux seems to be limited at hyperpolarised membrane potentials. An essentially similar relationship was observed when GABA was added to the bath solution, and the increase in the membrane current at various membrane potentials measured using the voltage-clamp technique.

Since the time course of changes in i.p.s.c. became faster on hyperpolarisation of the membrane, nonlinearity may be explained by the voltage-dependent rate constants for the opening and closing of the conductance channels[9]. In many cases, however, the current–voltage relationship deviated more markedly from linearity than would be expected from the dependence of rate constants on voltage. Qualitatively similar nonlinearity—by the depolarisation of the membrane—has been observed at acetylcholine receptors in the frog endplate[7] and electroplaque[10]; these could be accounted for using the above hypothesis.

K. ONODERA
A. TAKEUCHI

Department of Physiology,
School of Medicine,
Junendo University,
Hongo, Tokyo 113, Japan

1 Takeuchi, A., and Takeuchi, N., *J. Neurophysiol.*, **22**, 395–411 (1959).
2 Gage, P. W., and Moore, J. W., *Science*, **166**, 510–512 (1969).
3 Kordaš, M., *J. Physiol., Lond.*, **204**, 493–502 (1969).
4 Magleby, K. L., and Stevens, C. F., *J. Physiol., Lond.*, **223**, 151–171 (1972).
5 Llinás, R., Joyner, R. W., and Nicholson, C., *J. gen. Physiol.*, **64**, 510–535 (1974).
6 Onodera, K., and Takeuchi, A., *J. Physiol., Lond.*, **252**, 295–318 (1975).
7 Dionne, V. E., and Stevens, C. F., *J. Physiol., Lond.*, **251**, 245–270 (1975).
8 Boistel, L., and Fatt, P., *J. Physiol., Lond.*, **144**, 176–191 (1958).
9 Magleby, K. L., and Stevens, C. F., *J. Physiol., Lond.*, **223**, 173–197 (1972).
10 Lassignal, N. L., and Martin, A. R., *Science*, **191**, 464–466 (1976).

α-adrenergic receptors and pacemaker current in cardiac Purkinje fibres

In recent years, membrane phenomena in cardiac fibres have been analysed with increasing accuracy, by the use of electronic feedback. Ionic transmembrane currents have now been measured and analysed for their rectifier properties, their time constants and their activation ranges. One of the current components analysed in much detail is the pacemaker current in cardiac Purkinje fibres. Evidence is now accumulating that this and other current systems are related to adrenergic receptors. Strong evidence has been put forward that this current is dependent on β-adrenergic receptors. In this paper we report results of tests carried out to investigate the involvement of α-adrenoceptors in controlling this process and to show that they are not directly involved.

The pacemaker potential in cardiac Purkinje fibres is generated by a time-dependent deactivation of a slow outward current, i_{K2}[1,2], which allows the time-independent steady-state inward current to depolarise the membrane. Using the formalism of Hodgkin and Huxley[3] this current component is described in the following way

$$i_{K2} = \overline{i_{K2}} \times s \qquad (1)$$

where $\overline{i_{K2}}$ is the instantaneous or fully activated current–voltage relationship and s is a dimensionless variable which controls the degree of activation of i_{K2} and which follows first-order kinetics

$$ds/dt = \alpha_s(1-s) - \beta_s \times s \qquad (2)$$

α_s and β_s are the rate constants which are voltage dependent only. The steady-state degree of activation of i_{K2} was measured in voltage clamp conditions[2,4] as the amplitude of the current tail on the return to the holding potential. Since the holding potential is unchanged, $\overline{i_{K2}}$ must be the same and equation (1) can be modified

$$i_{K2} \propto s \qquad (3)$$

Adrenaline shifts the S-shaped curve relating the steady-state degree of activation and the membrane potential (s_∞) as well as the reciprocal of the time constant ($\tau^{-1} = \alpha_s + \beta_s$) in the depolarising direction[5-9]. The instantaneous current–voltage relation,
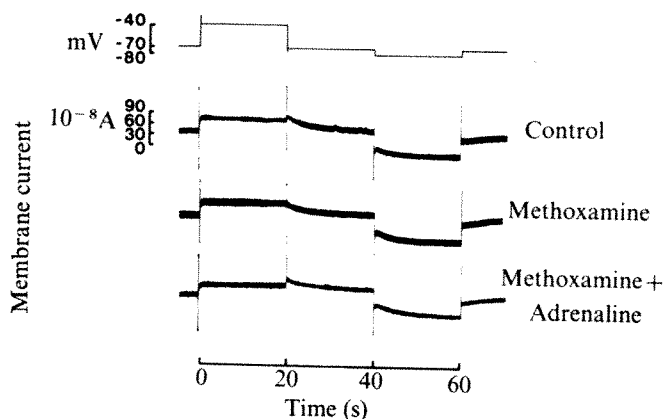
Fig. 1 Effect of methoxamine $(4 \times 10^{-5}\,\mathrm{g\,ml^{-1}})$ on voltage clamp currents in sheep Purkinje fibres. The Tyrode solution contained the following concentrations of ions (mM): Na^+ 148.3, K^+ 4, Ca^{2+} 1.8, Mg^{2+} 0.5, Cl^- 144.6, HCO_3^- 12, PO_4^{3-} 0.35, glucose $2\,\mathrm{g\,l^{-1}}$. Each of the current records shows the response to 30 mV depolarisation and to 10 mV hyperpolarisation from a holding potential of $-70\,\mathrm{mV}$. Note that the time-dependent current (i_{K2}) is changed neither with methoxamine nor with methoxamine plus adrenaline. The instantaneous current jump following depolarisation vanishes with methoxamine and is restored by adding adrenaline in the presence of methoxamine.

however, is altered neither by adrenaline[6,8], nor by racemates[8], nor optical isomers[9] of various β-receptor blockers. It was previously thought that in heart muscle only β-adrenergic receptors are present and thus it was rather surprising when Cranefield et al.[10] reported evidence for α-receptors in Purkinje fibres. Since agents specifically blocking β-adrenergic receptors shift the kinetics of i_{K2} back in the negative direction[5-9] the question arose whether i_{K2} may also be influenced by drugs stimulating alpha receptors. Tsien[6] found no detectable shift of the s-kinetics using phenylephrine $(10^{-5}\,M)$ and propranolol $(10^{-6}\,M)$, a well known β-receptor-blocking agent with strong local anaesthetic properties. Since in this (single) experiment two drugs were used—one of them (propranolol) with appreciable side-effects—it seemed appropriate to extend these studies and to use methoxamine $(4 \times 10^{-5}\,\mathrm{g\,ml^{-1}})$[10], a very selective α-receptor-stimulating agent with β-receptor-blocking activity[11] (provided by the Deutsche Wellcome GmbH).
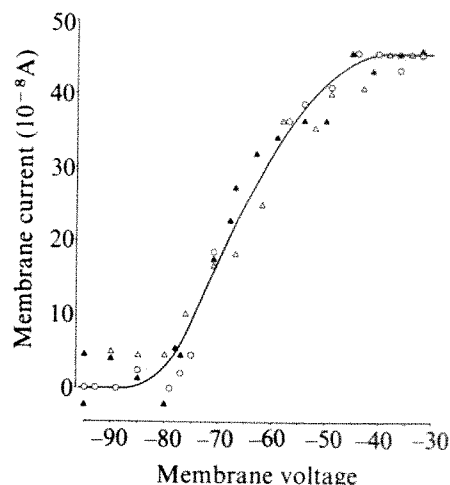
Fig. 2 Relationship between the steady-state degree of activation (s_∞) of i_{K2} and the membrane potential. ○, Control; △, methoxamine $(4 \times 10^{-5}\,\mathrm{g\,ml^{-1}})$; ▲, adrenaline $(10^{-6}\,\mathrm{g\,ml^{-1}})$ plus methoxamine. Note that unlike the treatment with β-receptor-stimulating agents methoxamine as an α-adrenergic substance does not change the voltage dependence of s_∞. Nor does adrenaline cause any shift in the presence of methoxamine.

In our experiments chart records of Purkinje fibre action potentials showed that a fibre was slightly depolarised with methoxamine. For an example in control conditions the resting potential amounted to -76 mV and decreased during the administration of methoxamine to -72 mV. Thus, it is not possible to tell whether the diminution of the amplitude of the action potential (and probably also of the rate of rise) observed after treatment with methoxamine was due to depolarisation only or whether this drug also reduces the availability of the excitatory sodium current at a particular membrane potential.

Weidmann's[12] h_∞ curve predicts a reduction in the rate of rise by about 25% due to a depolarisation of the order we observed. But in spite of the slight depolarisation, the time course of the pacemaker potential was essentially unchanged.

Figure 1 shows the membrane current changes as responses to long-lasting depolarising and hyperpolarising clamp pulses. There was no measurable change in the time-dependent currents. The only alteration was in the non-time-dependent background current called i_{K1} (refs 2 and 13). This instantaneous current changed in the inward direction—that is, the outward current jump after depolarisation became negligible when methoxamine was present; during and after hyperpolarisation an increase in i_{K1} could hardly be detected. But a decrease in an instantaneous outward current would be expected to depolarise the fibre. This current change was reversed when adrenaline was added together with methoxamine.

Figure 2 shows the relationship between the steady-state
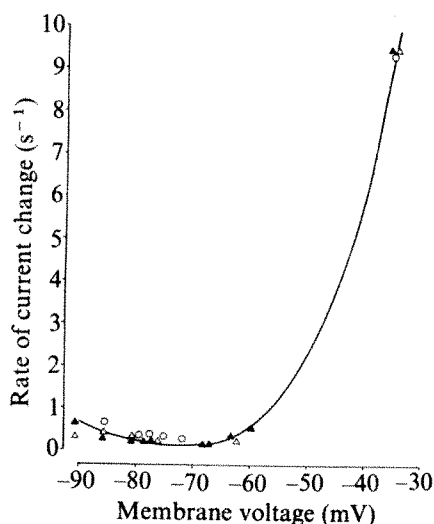
Fig. 3 Voltage dependence of the rate of change of i_{K_2} measured in terms of reciprocal of the time constant (τ^{-1}) of current change. Symbols as in Fig. 2.

degree of activation of i_{K_2} and the membrane potential measured as the amplitude of the current tail on return from various clamp potentials to the holding potential. This curve was shifted neither with methoxamine nor with methoxamine plus adrenaline (Fig. 2).

The absence of any shift in s_∞ by adrenaline in the presence of methoxamine was probably a consequence of the β-receptor-blocking activity of this drug[11].

The time constant of i_{K_2} at potentials less negative than the pacemaker region of potentials is determined by a procedure called the 'envelope test'[2] in which the duration of the clamp to a particular potential is progressively shortened, thereby switching on i_{K_2} to various degrees.

The time constant of i_{K_2} was identical at a particular potential whether determined in control conditions, in the presence of methoxamine or in the presence of methoxamine plus adrenaline. τ_s was measured as the amplitude of current tails on return to the holding potential following clamps to -36 mV of various durations and was found to be 106 ms. Nor is there any shift in the U-shaped curve relating the reciprocal of the time constant and the membrane potential (Fig. 3).

The kinetics of i_{K_2}—or the second factor in equation (1)—are related to the electric field at the membrane and thus the voltage dependence of the kinetics may be modified by altering the amount of electric charges at the inner or outer edge of the

Fig. 4 Instantaneous current–voltage relationship of i_{K_2}. This relationship is obtained from the quotient of the amplitudes of i_{K_2} during and following voltage clamp pulses to various levels. The 'rectifier ratio' obtained in this way is then multiplied by the amplitude of s_∞ measured at the same holding potential, resulting in the value of the 'rectifier function', $\overline{i_{K_2}}$. The absolute magnitude of i_{K_2} is altered neither with methoxamine nor with methoxamine plus adrenaline and the rectifier properties are unchanged with either condition. Symbols as in Figs 2 and 3.

Rectifier function
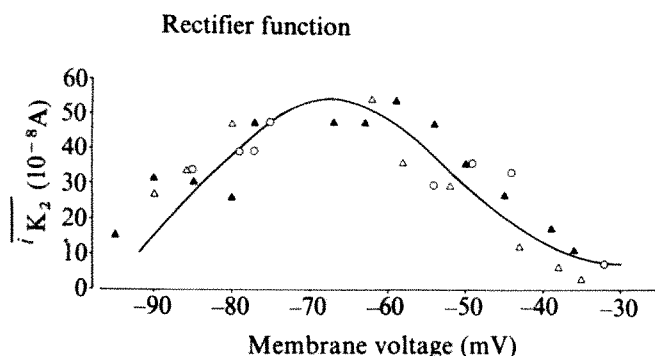
membrane. The findings described here show that stimulation of α-receptors does not influence the electric field near the i_{K_2} channel.

The instantaneous current voltage relation of i_{K_2} usually shows a marked inwardly directed rectification and a considerable negative slope[2].

The results reported here (Fig. 4) demonstrate that neither methoxamine nor methoxamine plus adrenaline alter the absolute magnitude or the rectifier properties of i_{K_2}.

Our results support the idea that i_{K_2} is governed by β-receptors[6-9] and they exclude the possibility that α-receptors influence i_{K_2}. Thus, it is not surprising that drugs selectively stimulating α-receptors do not have a positive chronotropic action on cardiac muscle.

OTTO HAUSWIRTH
HEINZ D. WEHNER
ROLF ZISKOVEN

Department of Physiology II,
University of Bonn,
D53 Bonn, West Germany

1 Vassalle, M., *Am. J. Physiol.*, **210**, 1335 (1966).
2 Noble, D., and Tsien, R. W., *J. Physiol., Lond.*, **195**, 185 (1968).
3 Hodgkin, A. L., and Huxley, A. F., *J. Physiol., Lond.*, **117**, 500 (1952).
4 Deck, K. A., Kern, R., and Trautwein, W., *Pflügers Arch. ges. Physiol.*, **280**, 50 (1964).
5 Hauswirth, O., Noble, D., and Tsien, R. W., *Science*, **162**, 916 (1968).
6 Tsien, R. W., *J. gen. Physiol.*, **64**, 293 (1974).
7 Tsien, R. W., *J. gen. Physiol.*, **64**, 320 (1974).
8 Hashimoto, K., Hauswirth, O., and Ziskoven, R., *Pflügers Arch. ges Physiol.*, **359**, R30 (1975).
9 Hauswirth, O., Wehner, H. D., and Ziskoven, R., *Pflügers Arch. ges. Physiol.* (in the press).
10 Cranefield, P. F., Hoffman, B. F., and Wit, A. L., *Nature new Biol.*, **234**, 159 (1971).
11 Imai, S., Shigei, T., and Hashimoto, K., *Circulation Res.*, **9**, 552, (1961).
12 Weidmann, S., *J. Physiol., Lond.*, **127**, 213 (1955).
13 Noble, D., *J. Physiol., Lond.*, **160**, 317 (1962).

Spontaneous repetitive hyperpolarisations from cells in the rat adenohypophysis

SPONTANEOUS, repetitive potential changes in the form of action potentials have been observed in only a few types of highly specialised cells such as cardiac pacemaker cells, various smooth muscle cells and certain nerve cells. Some endocrine cells, however, can exhibit repetitive depolarisations with similarities to action potentials; for example, the beta cells of the pancreas[1], cells in the adrenal cortex[2] and cells of a clonal line of an anterior pituitary tumour[3]. We report here that some cells in the adenohypophysis of ovariectomised rats are electrically excitable as well as exhibiting spontaneous, repetitive hyperpolarisations.

The rats were ovariectomised 3–5 weeks before the experiments, to induce hypertrophy of the pituitary gonadotropic cells and produce a more uniform population of cells from which more stable recordings of membrane potentials could be obtained. One hemilobe of the anterior pituitary gland was mounted in a Perspex chamber and superfused with Krebs–Henseleit bicarbonate solution at 37 °C, in a manner similar to that described before[1]. Membrane potentials were recorded from superficial cells by means of glass microelectrodes filled with 5 M potassium acetate with a resistance of 70–150 MΩ. A bridge circuit was used for intracellular recording and injection of current.

It was difficult to obtain stable measurements of membrane potentials from the cells of the adenohypophysis, but in 13
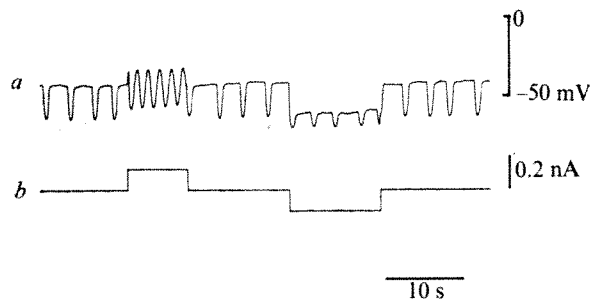
Fig. 1 Intracellular recording from an anterior pituitary cell exhibiting spontaneous, repetitive hyperpolarisation. The cell, which had a resting membrane potential of -42 mV, exhibited spontaneous hyperpolarisations at a frequency of 25 min^{-1} for 4 min before the record in a was obtained. The left part shows the effect of a depolarising current, and the right shows the effect of a hyperpolarising current. b Indicates that the current passed through the microelectrode. This and all subsequent records, were made by means of a pen recorder and the pituitary was obtained from a rat ovariectomised 4 weeks before study.

successful experiments recording from 127 cells the mean value was -49.3 ± 2.6 mV (s.e. of mean), which is more negative than previous *in vivo* measurements[4] (-21 mV) and *in vitro* measurements on pituitary tumour cells[3] (-41 mV). In eight experiments in which the input resistance—taken as a measure of the cell membrane resistance—was determined by passing small depolarising currents with the bridge balanced before and after penetration, the mean value was 118 MΩ (range 5–330 MΩ). This value is slightly larger than that found in a previous study[4] of *in vivo* rat pituitary (42 MΩ) but less than the value observed[3] with a clonal line of pituitary tumour cells (340–700 MΩ).

One population of cells with a mean resting potential of -37.1 ± 1.7 mV ($n=12$) and a mean resistance of 111 ± 31 MΩ ($n=7$) was peculiar in exhibiting spontaneous, repetitive hyperpolarisations. The mean frequency was 37.5 min^{-1} (range 2–65) and the mean amplitude was 16.0 mV (range 8–26). The record shown in Fig. 1 demonstrates the phenomenon and shows that the frequency of the hyperpolarisations increased when a small depolarising current was passed through the membrane. When the cell was hyperpolarised by a current passed in the opposite direction, the amplitude of the hyperpolarisations was reduced substantially, whereas the frequency

Fig. 2 Intracellular recordings demonstrating the excitability of the hyperpolarising cells. a, Ability of a depolarising current (indicated on the lower trace) and b, ability of termination of a hyperpolarising current (indicated on the lower trace) to evoke repetitive hyperpolarisations from an otherwise inactive cell.
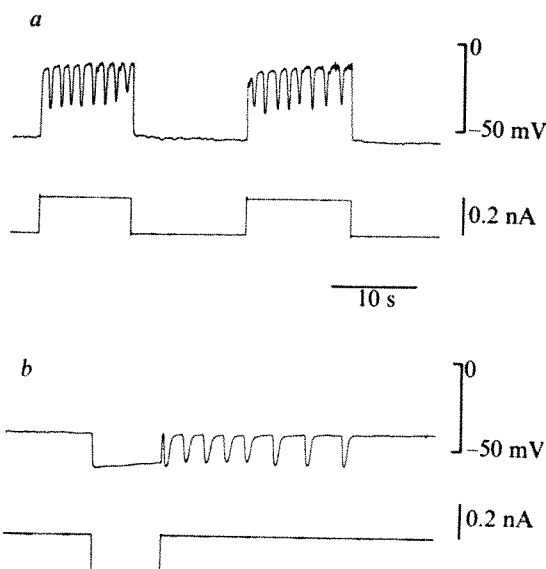
did not change appreciably. Although most of these cells continued their repetitive hyperpolarisations for as long as 24 min, in a few cases the hyperpolarisations ceased even though the resting potential remained unchanged. In these cells, as well as in some cells which did not exhibit spontaneous hyperpolarisations, it was possible to initiate repetitive hyperpolarisations either by passing a depolarising current or by termination of a hyperpolarising current (Fig. 2).

A few cells transiently increased their resting potentials to values as negative as -75 mV. When this happened, the magnitude of the repetitive hyperpolarisations decreased successively and they were abolished at membrane potentials ranging from -55 to -70 mV. When the resting potential exceeded the value at which the hyperpolarisations were abolished, the polarity of the potential changes was reversed. Subsequently the resting potential of these cells returned to the original value (close to -40 mV) whereby the hyperpolarising behaviour was restored. It was also possible to change the spontaneous, repetitive hyperpolarisations to depolarisations by passing a hyperpolarising current sufficient to increase the resting potential beyond the 'reversal potential' (Fig. 3a).

Since hyperpolarisation of cell membranes is frequently a result of an increased permeability to potassium ions, we
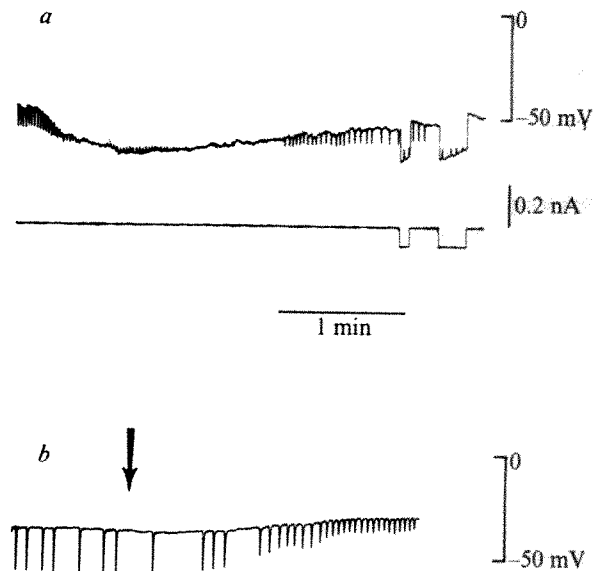
Fig. 3 Intracellular recordings from spontaneously, repetitively hyperpolarising cells. a, Changes in amplitude and polarity of the spontaneous, repetitive potential changes caused by variations in the resting membrane potential. The left part shows the effect of a spontaneous increase in the resting potential, and the right shows the effect of hyperpolarisation caused by a current (indicated on the lower trace) passed through the microelectrode. b, Effect of increasing the extracellular potassium ion concentration on amplitude and frequency of repetitive hyperpolarisations. From the time indicated by the arrow, a modified Krebs–Henseleit bicarbonate solution with a potassium ion concentration of 47 mM (10 times normal, sodium ion concentration reduced correspondingly) was passed through the superfusion chamber at a rate of 0.78 chamber volume per min.

studied the effect of raising the potassium ion concentration of the superfusion fluid tenfold (to 47 mM). Figure 3b demonstrates the resulting decrease in the amplitude of the hyperpolarisations; the hyperpolarisations were subsequently abolished (not shown). The resting membrane potential was reduced by the high potassium ion concentrations, and this reduction was associated with an increase in the frequency of the hyperpolarisations. All the effects of the high extracellular potassium ion concentration were reversible.

Although it is not known from which cells in the adenohypophysis the spontaneous, repetitive hyperpolarisations originate, it is possible that the potentials have been recorded from

the gonadotropic cells, which exhibit selective hypertrophy after ovariectomy. Our inability to record spontaneous, repetitive hyperpolarisations from the adenohypophyses of normal rats gives some support to this view.

In cells able to hyperpolarise in response to stimulation, for example, nerve cells stimulated by inhibitory transmitter substances[5], salivary gland acinar cells stimulated by autonomic nerves[6] and certain photoreceptor cells stimulated by light[7], a common underlying mechanism is a stimulation-induced increase of the conductance of the cell membranes to potassium ions. In contrast, the cells in the adenohypophysis can produce repetitive hyperpolarisations in the absence of any obvious discontinuous, repetitive stimulus. The effect of increasing the extracellular potassium ion concentration indicates that a cyclic, transient increase in the conductance of the cell membrane to this ion may be involved in the hyperpolarisations. On the other hand, the value of the reversal potential is compatible with additional conductance changes to other ions such as sodium and calcium. The latter possibility is interesting because uptake of extracellular calcium ions is an important step in many secretory processes[8].

Although our results suggest that the hyperpolarising cells have pacemaker-like properties, there is an alternative possibility. If the hyperpolarising cells were connected to spontaneously discharging nerve cells through both an inhibitory chemical and an electrical synapse, the observed behaviour could be explained by postulating that the electrical synapse showed rectification. Neurones connected by both electrical and chemical synapses have been demonstrated[9,10]; neurones innervating adenohypophyseal secretory cells, however, have not been demonstrated.

The excitability of the hyperpolarising cells (Figs 1 and 2) show certain similarities to the excitability found in depolarising pituitary tumour cells by Kidokoro[3]. The cells described here however, are unique both in hyperpolarising in response to direct electrical stimulation and in showing repetitive hyperpolarisations.

This work was supported by the NIGMS of the USPHS, the Cystic Fibrosis Foundation, the Danish Medical Research Council and the NOVO Foundation.

JORGEN HEDEMARK POULSEN*
JOHN A. WILLIAMS

Department of Physiology,
University of California, San Francisco,
California 94143

1 Dean, P. M., and Matthews, E. K., *J. Physiol., Lond.*, **210**, 255–264 (1970).
2 Matthews, E. K., and Saffran, M., *Nature*, **219**, 1369–1370 (1968).
3 Kidokoro, Y., *Nature*, **258**, 741–742 (1975).
4 York, D. H., Baker, F. L., and Kraicer, J., *Neuroendocrinology*, **8**, 10–16 (1971).
  York, D. H., Baker, F. L., and Kraicer, J., *Neuroendocrinology*, **11**, 212–228 (1973).
5 Eccles, J. C., *The Physiology of Synapses*, 173–188 (Springer, Berlin, Gottingen and Heidelberg and Academic, New York, 1964).
6 Nishiyama, A., and Petersen, O. H., in *Secretory Mechanisms of Exocrine Glands* (edit. by Thorn, N. A., and Petersen, O. H.,) 216–224 (Munksgaard, Copenhagen, 1974).
7 McReynolds, J. S., and Gorman, A. L. F., *Science*, **183**, 658–659 (1974).
8 Douglas, W. W., *Br. J. Pharmac.*, **34**, 451–474 (1968).
9 Martin, A. R., and Pilar, G., *J. Physiol., Lond.*, **171**, 454–475 (1964).
10 Landmesser, L., and Pilar, G., *J. Physiol., Lond.*, **222**, 691–713 (1972).

Increase in microviscosity with ageing in protoplast plasmalemma of rose petals

IT is now widely accepted that most physiological functions of biological membranes are related to the dynamic characteristics of transport, enzyme and receptor sites. Since the rotational and translational mobility[1], as well as the degree of exposure of membrane proteins[2], are largely determined by the fluidity of the lipid layer, membrane microviscosity may be important in regulating various physiological activities, as in lymphocytes[3-5] and erythrocytes[6]. To our knowledge, changes in fluidity of plant plasmalemma, the outer surface of the protoplast[7], have not yet been reported. We report here a gradual increase in microviscosity of protoplast plasmalemma from petals of ageing rose flowers.

Rose (*Rosa* hyb.) flowers cv. Golden Wave (Syn. Dr. Verhage) were allowed to develop and age on the plant and were picked at different stages of development—from a tight bud to a fully opened flower with first signs of wilt and colour fading. The four outer petals of the flowers were macerated enzymatically in 0.6 M mannitol (Merck) containing 0.5% cellulysin (Calbiochem) and 0.25% driselase (Kyowa Hakko Koggo Co.) by incubation for 18 h at 20 °C in the dark. As a reporter molecule, we used 1,6-diphenyl-1,3,5-hexatriene (DPH), the most efficient probe available for monitoring fluidity properties of lipid regions by fluorescence polarisation[2-4,8]. For labelling, isolated protoplasts were incubated for 1 h at 37 °C with 10^{-6} M dispersion of DPH in 0.6 M mannitol. The labelled protoplasts were then washed and resuspended in 0.6 M mannitol to a concentration of 5×10^6–8×10^6 ml^{-1}. A fluorescence microscopic view of a DPH-labelled protoplast is of a glowing periphery, indicating that the fluorescence signal is generated predominantly from the hydrocarbon layer of the protoplast plasmalemma.

Fluorescence polarisation and intensity in the labelled protoplast suspension were recorded, and the corresponding microviscosity evaluated by a method described previously[2,3,8]. Figure 1 shows the change in microviscosity of the protoplast plasmalemma with stage of development of the petals. As shown, the membrane microviscosity ($\bar{\eta}$) gradually increases with the flower development. The rate of increase in $\bar{\eta}$ is markedly enhanced during the 4-d period between stage C (partially opened flower) and stage E (wilting flower), when the microviscosity is approximately doubled. The effect of temperature on the plasmalemma microviscosity was also determined in each stage in the temperature range of 4–30 °C. For all membranes, the change of log $\bar{\eta}$ with the reciprocal of the absolute temperature ($1/T$) followed a straight line with an apparent flow activation energy (ΔE) of 3.4 ± 0.2 kcalorie mol^{-1}. As an

Fig. 1 Microviscosity of protoplast plasmalemmas obtained from rose petals at different stages of development. Bars represent the range of values obtained with protoplasts from 3–5 different flowers.

6 Wiley, J., and Cooper, R. A., *Biochim. biophys. Acta*, **413**, 425–431 (1975).
7 Ruesink, A. W., *Pl. Physiol.*, **47**, 192–195 (1971).
8 Shinitzky, M., and Barenholz, Y., *J. biol., Chem.*, **249**, 2652–2658 (1974).
9 Lösel, D. M., and Lewis, D. H., *New Phytol.*, **73**, 1157–1169 (1974).
10 Newman, D. W., Rowell, B. W., and Byrd, K., *Pl. Physiol.*, **51**, 229–233 (1973).
11 Mayak, S., Halevy, A. H., and Katz, M., *Physiol. Plant.*, **27**, 1–4 (1972).
12 Sacher, J. A., *A. Rev. Pl. Physiol.*, **24**, 197–224 (1973).
13 Borochov, A., Mayak, S., and Halevy, A. H., *Physiol. Plant.*, **36**, 221–224 (1976).

Fig. 2 Temperature dependence of microviscosity in protoplast plasmalemmas from rose petals at stages D (△) and E (○). Each point represents an average of three values obtained with different preparations.

example, Fig. 2 shows plots of log $\bar{\eta}$ against $1/T$ obtained for plasmalemma from stages D and E (see Fig. 1).

The data now available for plant membranes are still few and insufficient. The conclusions that can be drawn from this study are therefore bound to facts which were established for mammalian membranes. The unusually low and invariable value of $\Delta E = 3.4$ kcalorie mol^{-1} (in mammalian membranes $\Delta E = 6.5$–8.5 kcalorie mol^{-1}, (see ref. 2)) indicates that the lipid domains of the plasmalemma are highly ordered, even though they maintain a relatively high fluidity. In addition, it suggests that the membrane function and organisation are only slightly affected by temperature changes, which may be a special feature of plant membranes. A combination of these properties has no parallel in mammalian membranes, and indicates that the lipid composition, organisation and dynamics of plant and mammalian membranes are basically different. Nevertheless, the increase of plasmalemma microviscosity with the age of the flower could result primarily from one (or a combination) of the following changes: (1) an increase in the molar ratio of sterols to phospholipids; (2) a decrease in the degree of unsaturation of the fatty acid chains; (3) an increase in the relative amount of membrane proteins. Changes of the first[9] and the second[10] kind have been observed in aged plant tissues. The increase in microviscosity caused by such changes may decrease the turnover of membrane enzymes and hormone receptors and thus attenuate some metabolic processes which will eventually lead to the death of the flower[11–13].

A. BOROCHOV
A. H. HALEVY

Department of Ornamental Horticulture,
The Hebrew University, Rehovot, Israel

M. SHINITZKY

Department of Membrane Research
The Weizmann Institute of Science, Rehovot, Israel

1 Edidin, M., *A. Rev. Biophys. Bioengng*, **3**, 179–201 (1974).
2 Shinitzky, M., and Inbar, M., *Biochim. biophys. Acta*, **433**, 133–149 (1976).
3 Shinitzky, M., and Inbar, M., *J. molec. Biol.*, **85**, 603–615 (1974).
4 Inbar, M., and Shinitzky, M., *Eur. J. Immun.*, **5**, 166–170 (1975).
5 Alderson, J. C. E., and Green, C., *FEBS Lett.*, **52**, 208–211 (1975).

Spin label study of erythrocyte membrane fluidity in myotonic and Duchenne muscular dystrophy and congenital myotonia

INTACT erythrocyte membranes from patients with myotonic muscular dystrophy (MMD) have been shown by spin labelling (review of method in ref. 1) to have greater membrane surface fluidity than normal[2,3]. We have now evaluated the specificity of this phenomenon. We used erythrocytes from patients with MMD, and also from patients with Duchenne muscular dystrophy (DMD), as a model of dystrophy with no myotonia, and from patients with congenital myotonia (CM) as a model of myotonia without dystrophy.

MMD has been considered a defect of muscle surface membrane based on the persistence of repetitive membrane depolarisation after nerve block or neuromuscular blockade[4]. A similar locus has been postulated for the membrane defect is congenital myotonia[4]. Our previous biochemical and biophysical studies have suggested that in MMD, there are membrane abnormalities in erythrocytes as well as muscle membranes[2,3,5–8]. Membrane protein phosphorylation is lower than normal[5,6] and the stoichiometry of the sodium pump is unusual[7]. An abnormally large number of stomatocytes was noted by scanning electron microscopy in MMD as well as several other inherited muscle diseases, including CM and DMD[8]. Studies with stearic acid methyl ester spin labels have demonstrated that MMD erythrocyte membranes are more fluid near the membrane surface than are those of normal controls[2,3]. The data now reported indicate similarly high erythrocyte fluidity near the membrane surface in CM, whereas erythrocytes from DMD and other non-myotonic muscular dystrophic patients have normal membrane fluidity as assessed using 5-NMS (see below). These results suggest a correlation of increased membrane fluidity with the presence of myotonia.

The principal spin label we used was 2-(3-carboxypropyl)-2-tridecyl-4,4-dimethyl-3-oxazolindinyloxyl methyl ester (5-nitroxide methyl stearate) (5-NMS) (Syva Associates). The erythrocytes were obtained from fresh heparinised blood from patients and from age and sex-matched controls. In experiments with MMD cells, data previously obtained[3] from seven normal and nine myotonic patients were recalculated as described below. Samples from twelve normal and ten DMD patients and nine normal and nine CM patients, respectively, were used. In each of the three separate sets of experiments, intact erythrocytes were prepared by centrifuging blood at 1,570g for 10 min at 4 °C and washing the cells four times with 5 mM sodium phosphate–150 mM NaCl buffer, pH 8.0 (PBS). The buffy coat was removed carefully and the cells were resuspended in isotonic PBS at a haematocrit of 45–50%.

Intact MMD cells were labelled overnight at 37 °C and intact CM and DMD cells were incubated with 5-NMS at room temperature for 10–30 min. In each respective set of experiments, control and disease state membranes were treated identically. The ratio of spin label to lipid molecules were kept at approximately 1:50 to avoid spin exchange effects[9]. Magnetic resonance measurements were performed as before[3] on both a Varian V-4502-15 and a Ventron-Magnion MVR-9X electron spin resonance (ESR) spectrometer, with a quartz aqueous sample cell.

The spin label we used is thought to orient in the erythrocyte membrane with the polar head group near the lipid head group and the alkyl chain parallel on the average to the lipid alkyl

chains. Rapid anisotropic rotation takes place about the long axis of the spin label whose nitroxide group probes an environment near the membrane surface. A typical spectrum of 5-NMS in erythrocyte membranes is shown in Fig. 1, in which the experimental T-tensor parameters are indicated. The order parameter, S, is a measure of the fluidity of the local environment in which the paramagnetic centre of the spin label is found while a_N, the nitrogen isotropic hyperfine coupling constant, is a probe of local polarity. Regardless of whether the random walk model of Jost *et al.*[10] or the model of Mason *et al.*[11] (which involves an inverse relationship between the rate of rotation of the symmetry axis and the apparent values of both S and a_N) is used to explain spectra like those of Fig. 1, the interpretation of S as a measure of local fluidity is valid. The polarity correction factor used by Hubbell and McConnell[12] with the crystal T-tensor values of "doxyl"-propane reported by Jost *et al.*[10], were used in the calculation of S.
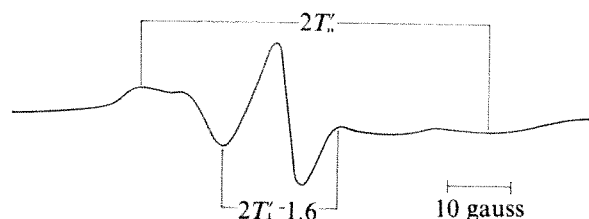
Fig. 1 Typical spectrum of 5-NMS in normal erythrocytes. The tensor parameters are indicated.

The mean S values for controls and each disease state performed in the three separate sets of experiments were compared by a two-way analysis of variance[13] (Table 1). This two-tailed method of statistical analysis minimises possible fluctuations from day to day between separate experimental values which can often occur with biological samples.

P is the significance of the difference of the mean values of $S_{control}$ and $S_{disease\ state}$ (or $(a_N)_{control}$ and $(a_N)_{disease\ state}$) in each set of experiments calculated using a two-way analysis of variance. The null hypothesis was that MMD, DMD and CM membranes are, respectively, not different from the corresponding controls. The difference of $S_{control}$ and S_{MMD} or S_{CM} is significantly greater than zero, indicating more fluid environments in erythrocytes from these myotonic conditions. In contrast, there is no mean difference of $S_{control}$ and S_{DMD} and, accordingly, no difference in surface membrane fluidity. In this latter case the data actually tended to show DMD erythrocytes to be more rigid than controls but statistical significance of this observation could not be demonstrated.

The difference between the mean values of the control parameters in the MMD experiments compared with those in the CM and DMD experiments probably reflects the different

Table 1 Comparison of the order parameter S for 5-NMS in normal and MMD, DMD or CM intact erythrocytes

	S_N	S_{MMD}	S_N	S_{DMD}	S_N	S_{CM}
Mean	0.625	0.590	0.598	0.605	0.588	0.570
s.e.m.	0.008	0.006	0.012	0.015	0.010	0.008
n	7	9	12	10	9	9
P	< 0.005		$0.1 < P < 0.25$		< 0.01	

N, normal. S is calculated by

$$\frac{T_{\parallel}' - T_{\perp}' \quad a_{xl}}{T_{\parallel} - T_{\perp} \quad a'}$$

where the primed values are obtained experimentally (Fig. 1) and $a' = (T'_{\parallel} + 2T'_{\perp})/3$. The unprimed crystal values are obtained from the results of Jost *et al.*[10] on doxyl-propane. The data for MMD were from ref. 3, recalculated by a two-way analysis of variance. Data from three separate sets of experiments are presented. P was calculated by a two-way analysis of variance using a two-tailed test with a null hypothesis that the respective means are equal.

methods of labelling used. Intact erythrocytes held for several hours at 37 °C rapidly lose intracellular ATP with a consequent decrease in cell deformability[14]. This latter property may reflect an increased membrane rigidity which in turn would demonstrate an increased order parameter. Intact erythrocytes labelled overnight at 37 °C show an increased S-value when compared with cells from the same donor labelled for 15 min at room temperature (D.A.B., unpublished). The a_N value of 5-NMS in control cells is also increased in the MMD experiments compared with those of the other two sets of experiments (Table 2), suggesting that the spin probe is experiencing a different environment in the former case. The important aspect of these data in that within each separate set of experiments, normal and disease state membranes were treated in the same manner.

Our previous results[3] recalculated by the two-way analysis of variance show MMD erythrocyte membranes to have greater than normal surface fluidity as assessed by 5-NMS and less than normal local polarity, as measured by 16-NMS, relative to those of normal controls. The study reported here demonstrates that membrane fluidity is greater than normal in CM but not different from controls in DMD, hyperkalaemic periodic paralysis (without para-myotonia) or oculopharyngeal muscular dystrophy. There are no highly significant polarity changes as measured by a_N in the various experiments (Table 2).

We also found greater than normal membrane fluidity in MMD and CM erythrocytes when we used the corresponding acid spin label (5-NS) in identical procedures. The same differences of fluidity in erythrocyte ghost preparations[6] as in intact cells were also observed.

Table 2 Comparison of a_N for 5-NMS in normal and MMD, DMD or CM intact erythrocytes

	N	MMD	N	DMD	N	CM
Mean	15.91	15.81	15.56	15.67	15.50	15.43
s.e.m.	0.07	0.05	0.05	0.31	0.31	0.32
n	7	7	10	10	9	9
P	< 0.25		< 0.25		< 0.25	

The data for MMD were from ref. 3, recalculated by a two-way analysis of variance.
P was calculated as in Table 1.

Increased membrane fluidity has been demonstrated in both MMD and CM erythrocyte membranes by means of ESR spectroscopy. Although this method detects a fluidity increase in both cases, the specific biochemical alterations responsible for the change in these diseases may be quite different. Lower than normal phosphorylation of component a has been demonstrated in MMD erythrocyte ghosts, while CM ghosts do not have such biochemical alteration[16]. Physiological data also support different mechanisms of myotonia. Muscle membranes from CM demonstrate greater than normal resistance and less than normal chloride conductance, while these characteristics are not apparently affected in MMD[4]. It is interesting that the two myotonic diseases MMD and CM involve abnormal membrane fluidity of erythrocytes whereas DMD and the other two non-myotonic dystrophies apparently do not.

Membrane fluidity is a function of many variables, including the nature of membrane lipids (type of phospholipid, fatty acid chain length and degree of unsaturation, head group and cholesterol content), temperature, water content, presence of ions (especially divalent cations), and the total of protein-lipid interactions[16,17]. Activity of membrane-associated transport enzymes in bacterial[18] and animal[19,20] systems can be altered by affecting membrane fluidity. Increased membrane fluidity in MMD and CM may affect the expression of a transport or enzyme system resulting in the clinical sign of myotonia in muscle. Although the specific mechanism responsible for these

phenomena is unknown, our spin label measurements suggest a correlation of increased erythrocyte membrane fluidity with the presence of myotonia.

D. Allan Butterfield*

Department of Chemistry,
University of Kentucky,
Lexington, Kentucky 40506

D. B. Chesnut

Paul M. Gross Chemical Laboratory,
Duke University,
Durham, North Carolina 27706

Stanley H. Appel
Allen D. Roses

Division of Neurology,
Duke University School of Medicine,
Durham, North Carolina 27710

1 *Spin Labelling—Theory and Applications* (edit. by Berliner, L. J.) (Academic, New York, 1976).
2 Butterfield, D. A., Chesnut, D. B., Roses, A. D., and Appel, S. H., *Proc. natn. Acad. Sci. U.S.A.*, 71, 909–913 (1974).
3 Butterfield, D. A., Roses, A. D., Cooper, M. L., Appel, S. H., and Chesnut, D. B., *Biochemistry*, 13, 5078–5082 (1974).
4 Lipicky, R. J., and Bryant, S. H., in *New Developments in Electromyography and Clinical Neurophysiology*, 1 (edit. by Desmedt, J. E.), 451–463 (Karger, Basel, 1973).
5 Roses, A. D., and Appel, S. H., *Proc. natn. Acad. Sci. U.S.A.*, 70, 1855–1859 (1973).
6 Roses, A. D., and Appel, S. H., *J. Membr. Biol.*, 20, 51–58 (1975).
7 Hull, K., and Roses, A. D., *J. Physiol., Lond.*, 254, 169–181 (1975).
8 Miller, S. E., Roses, A. D., and Appel, S. H., *Archs Neurol.*, 33, 172–174 (1976).
9 Butterfield, D. A., Whisnant, C. C., and Chesnut, D. B., *Biochim. biophys. Acta*, 426, 697–702 (1976).
10 Jost, P. C., Libertini, L. J., Herbert, V. C., and Griffith, O. H., *J. molec. Biol.*, 59, 77–98 (1971).
11 Mason, R. P., Polnaszek, C. F., and Freed, J. H., *J. phys. Chem.*, 78, 1324–1329 (1974).
12 Hubbell, W. L., and McConnell, H. M., *J. Am. chem. Soc.*, 93, 314–326 (1971).
13 Brownlee, K. A., *Statistical Theory and Methodology in Science and Engineering* (Wiley, New York, 1960).
14 Weed, R. I., and LaCelle, P. L., in *Red Cell Membrane Structure and Function* (edit. by Greenwalt, T. J., and Jamieson, G. A.) (Lippincott, Philadelphia, 1969).
15 Roses, A. D., Appel, S. H., Butterfield, D. A., Miller, S. E., and Chesnut, D. B., *Trans. Am. neurol. Ass.*, 100, 17–20 (1976).
16 Jackson, R. L., and Gotto, A. M., *New Engl. J. Med.*, 290, 24–29 (1974).
17 Singer, S. J., *A. Rev. Biochem.*, 43, 805–833 (1974).
18 Overath, P., and Trauble, H., *Biochemistry*, 12, 2625–2634 (1973).
19 Grisham, D. M., and Barnett, R. E., *Biochemistry*, 12, 2635–2637 (1973).
20 Eletr, S., Zakim, D., and Vessey, D. A., *J. molec. Biol.*, 78, 351–362 (1973).

Vitamin D-stimulated intestinal calcium absorption may not involve calcium-binding protein directly

One of the earliest responses of chick intestine to cholecalciferol (vitamin D_3) is the appearance of a soluble calcium-binding protein (CaBP) for which current evidence indicates a role in the intestinal absorption of calcium[1,2]. CaBP, which has never been detected in the intestine of cholecalciferol-deficient chickens, is synthesised *de novo* in response to cholecalciferol, apparently by the induction of the transcription of a specific mRNA for CaBP (ref. 3). In the chick CaBP biosynthesis is initiated after a lag of about 7 h after physiological doses of the vitamin. A correlation between the appearance of CaBP and the first detectable increase in calcium absorption in response to cholecalciferol was observed when CaBP was measured using immunological procedures[4,5] but not using the less sensitive Chelex assay[6]. Consequently it has been generally accepted that CaBP plays a major role in chicks in cholecalciferol-stimulated calcium transport, but by a mechanism which has still to be described. These changes in intestinal

CaBP levels and in calcium absorption are brought about by 1,25-dihydroxycholecalciferol (1,25-$(OH)_2D_3$), the hormonal metabolite of D_3. We report here the sequence and time scale of changes in calcium absorption and CaBP levels in the intestine in response to 1,25-$(OH)_2D_3$. By comparison with events observed after cholecalciferol administration, the changes in response to the hormone turn out to be much more rapid and to have different rates of decay, with calcium transport being stimulated before CaBP can be detected immunologically.

Using immunoelectrophoresis[7] we have measured the amount of CaBP in the intestinal mucosa of rachitic chicks at various intervals after injecting 1,25-$(OH)_2D_3$ intracardially. CaBP was first detected 5 h after dosing (7 h earlier than after an equivalent dose of cholecalciferol) and by 16 h had increased to a level of 35 µg per g wet weight mucosa. Maximal quantities (42 µg g^{-1}) were obtained by 48 h, the levels declining to about 50% of the maximum by 72 h (Table 1); the protein could not be detected again by 96 h.

The ability of rachitic chicks to transport calcium across the intestine after receiving 1,25-$(OH)_2D_3$ was measured *in vitro* using everted intestinal sacs[8]. Sacs prepared from birds killed

Table 1 Calcium-binding protein levels in intestinal mucosa of rachitic chicks treated with 1,25-$(OH)_2D_3$

Time (h) after dosing	CaBP (µg per g wet weight tissue)	% Maximum response
0	0	0
5	*	
7	12.0	28.5
8	22.5	53.6
16	35.0	83.3
24	36.3	86.4
48	42.0	100
72	20.0	47.6
96	0	0

*, CaBP was detectable but in amounts too low for measurement. Groups of six rachitic chicks received intracardial injections of 125 ng 1,25$(OH)_2D_3$ in 0.1 ml 10% EtOH and 90% propylene glycol, and were killed at the times shown. CaBP was measured in the pooled duodenal cytosols by immunoelectrophoresis[7] using pure CaBP as standard.

2 h after dosing, transported calcium at a rate significantly higher than that of the controls, the maximum rate being reached by sacs obtained from birds 8 h after dosing (Fig. 1). By 21 h the ability of the sacs to transport calcium had declined to a rate slower than that occurring at 2 h. Thus by 4 h, 1,25-$(OH)_2D_3$-stimulated calcium transport was occurring at more than 50% of the maximum measured rate, whereas CaBP could not be detected. By 21 h, when calcium transport had declined markedly, the intestinal CaBP level was approaching its maximum. This lack of correlation between calcium transport and the quantity of CaBP induced by a single injection of 1,25-$(OH)_2D_3$, at both short and longer time intervals, suggested that the production of CaBP might be an event secondary to the hormonal stimulation of calcium transport. We therefore re-examined the phenomena using more sensitive methods for the detection of CaBP.

In the chick, cholecalciferol-stimulated CaBP appears to the greatest extent in the duodenum[9]. We therefore tested the ability of duodenal polysomes prepared from rachitic and 1,25-$(OH)_2D_3$-dosed chicks to synthesise CaBP in an *in vitro* cell-free system[11]. Incorporation of ^3H-leucine into completed polypeptide chains and into immunoprecipitable CaBP was determined. The specificity of the antiserum used to precipitate the tritiated CaBP was confirmed by analysing the dissociated immunoprecipitates on polyacrylamide gels. Typically, more than 80% of the total radioactivity recovered from the gel slices was localised in the CaBP peak. To measure CaBP synthesis, radioactivity in the CaBP area of the gels was related to the total TCA-precipitable incorporation into completed polypeptide chains.

Table 2 Effect of 1,25-(OH)₂D₃ on the synthesis of CaBP by polysomes *in vitro*

Time (h) after dosing	Radioactivity in immunoprecipitable CaBP (d.p.m.)	CaBP (% total proteins synthesised by polysomes)	% Maximum response
0	0	0	0
1*	0	0	0
2*	517	0.04	1.10
3	1,091	0.08	2.33
4	9,774	0.69	20.1
6	27,242	1.92	58.2
8	37,790	2.66	80.8
13	46,788	3.29	100
21	8,170	0.57	17.5
48	0	0	0

*Same birds were used for both polysome preparation and calcium transport measurements. At each other time, birds were taken from the same batch as those used for calcium transport.

Groups of four rachitic chicks received 125 ng 1,25-(OH)₂D₃ intracardially and were killed at the indicated times. Duodenal polysomes were prepared[10] and incubated for 40 min at 37 °C in triplicate 330-µl aliquots. The incubation medium contained 1 mM ATP, 0.1 mM GTP, 10 mM creatine phosphate, 40 µg ml⁻¹ creatine kinase, 50 µM of each amino acid except leucine, 5 µCi ml⁻¹ ³H-leucine, 0.5 mg ml⁻¹ polysomal RNA, 0.4 ml ml⁻¹ gel-filtered rat liver cell sap, 1 mM dithiothreitol, 4 mM MgCl₂, 100 mM KCl and 20 mM Tris-Cl pH 7.5. Duplicate 20 µl aliquots of the pooled triplicate postribosomal supernatants were removed for measurement of TCA-precipitable radioactivity in released protein. The remainder was immunoprecipitated with anti-CaBP as described previously[3]. The immunoprecipitate was dissolved by heating for 20 s at 90 °C in 100 µl SDS-Tris-dithiothreitol solution[12] and electrophoresed on SDS-acrylamide gels, which were then cut into 1-mm slices and processed for radioactive counting[3]. ³H radioactivity (d.p.m.) in the CaBP area of the gels is expressed as a percentage of the total TCA-precipitable incorporation into released proteins (corrected up to the maximum value obtained, that is, 1.4 × 10⁶ d.p.m. ml⁻¹) at each interval.

As shown in Table 2, CaBP mRNA activity was maximal 13 h after a single hormone injection, by which time CaBP represented 3–4% of the total protein-synthesising capacity of the polysomes (assuming a constant rate of translation). Polysomes from birds killed 1 h after receiving 1,25-(OH)₂D₃ were unable to synthesise CaBP. At this time, however, intestinal sacs prepared from the same birds could transport calcium at a low level (Fig. 1). Sacs prepared from birds killed 2 h after dosing, transported calcium at 27% of the maximum rate, but polysomes from these birds could only synthesise minute amounts of CaBP equivalent to 0.04% of the total proteins completed *in vitro*. Therefore, although translation
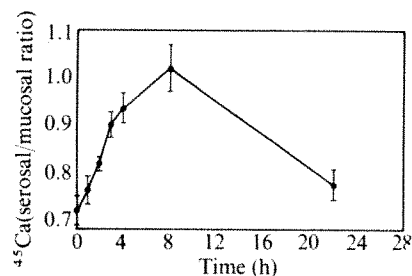
Fig. 1 Stimulation of *in vitro* calcium transport in chick intestine in response to 1,25-(OH)₂D₃. Rachitic chicks (other than controls) were dosed intracardially with 125 ng 1,25-(OH)₂D₃ and killed at the times shown. Sacs were constructed from 3.0 cm of ileum distal to the yolk sac remnant and filled with 0.3 ml medium. Incubation was carried out at 37 °C for 90 min with shaking in 25-ml Erlenmeyer flasks containing 3.0 ml of the same medium in 95% O₂/5% CO₂. A low sodium medium was used to reduce water transport[11]; this contained 198 mM mannitol, 25 mM KCl, 1.2 mM NaH₂PO₄, 25 mM NaHCO₃, 1.2 mM MgSO₄, 20 mM glucose, 0.5 mM CaCl₂ and ⁴⁵Ca at a concentration of 40,000 c.p.m. ml⁻¹. Aliquots (100 µl) were removed from the serosal and mucosal solutions after incubation for measurement of radioactivity. Each point is the mean of a group of 7–10 chicks. Vertical bars indicate ± s.e.m. A low level of calcium transport seemed to be taking place after 1 h (P < 0.3) but significant transport did not occur until 2 h (P < 0.001).

of CaBP mRNA *in vivo* had commenced by 2 h after dosing, CaBP production represented only 1.0% of the maximum activity attained.

Thus, at all times at which calcium transport could be observed at a significantly higher rate in dosed chicks than in rachitic birds, we detected nascent CaBP on the polysomes. We have also detected very low amounts (0.01% of total protein synthesis) of soluble (that is, completed, released from polysomes and presumably functional) ³H-CaBP synthesised *in vitro* by intestinal slices obtained from birds killed 2 h after
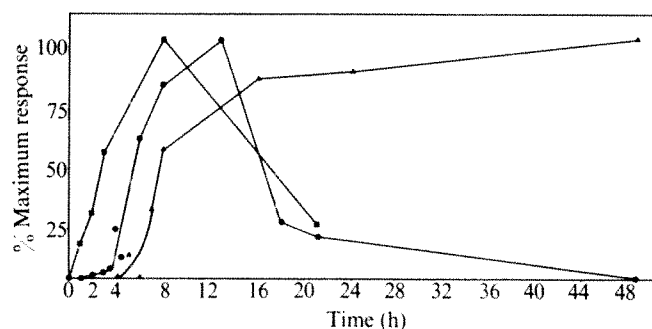
Fig. 2 Time course of events occurring in the rachitic chick intestine after a single intracardial injection of 125 ng 1,25-(OH)₂D₃. ■, Calcium transport *in vitro*; ●, CaBP mRNA activity; ▲, CaBP level in intestinal cytosol.

dosing (data not shown), confirming the presence of minute amounts of CaBP in the intestine at this stage. None was detectable at 1 h. From the composite data presented in Fig. 2, however, by 3 h after dosing, when calcium transport is proceeding at 50% of its maximum rate, CaBP biosynthesis is occurring at only 2.5% of its maximum rate, and in general the response curve of the latter lies about 2.5 h after the former.

It is also apparent from Table 2 and Fig. 2 that after 13 h the ability of the polysomes to synthesise CaBP declined rapidly, being completely absent by 48 h. This suggests that the half life of CaBP mRNA is not more than 4 h; high levels of CaBP observed in intestinal supernatants between 16 and 48 h and its disappearance between 48 and 96 h therefore indicate that the accumulated CaBP is stable within the 3–4-d life of the cell, rather than that a steady-state situation of continued synthesis and breakdown exists. In addition, the continued presence of CaBP in mucosal cells at times when calcium transport is no longer detectable indicates that other vitamin D-dependent factors are required.

There have been reports of a correlation between changes in calcium absorption and in CaBP levels which occur in chicks in response to cholecalciferol[4-6]. Before the present study, it seemed that a correlation could be demonstrated providing that the CaBP levels were measured by immunological procedures. Furthermore, the available information from two laboratories[13,14] on the effect of 1,25-(OH)₂D₃ on chick intestine seemed to support this view, but in both studies the times at which comparisons were made were too late to detect calcium transport in the absence of CaBP. In our studies, in which changes in calcium absorption and CaBP after 1,25-(OH)₂D₃ were compared at hourly intervals, calcium transport occurred while CaBP was absent or present in very small amounts. Thus, it is unlikely that CaBP is required for the initial calcium transport occurring in response to 1,25-(OH)₂D₃, although the possibility remains that only a small proportion (0.01–1.0%) of the CaBP ultimately synthesised in response to the hormone actively promotes calcium transport. If such a small proportion of CaBP is active in calcium absorption there are implications for the mechanism by which the protein is effective. That finding would imply a catalytic role for CaBP rather than its functioning in a stoichiometric relationship with the calcium being transported[15]. In addition, the intracellular distribution of this small quantity may be different from the distribution of the total CaBP eventually produced by the cell[16]. It seems

unlikely, however, that the bulk (> 99%) of the CaBP in the mucosal cell is inactive. We therefore conclude that there is not an obligatory requirement for CaBP in all calcium transport.

Whether such a conclusion is applicable to mammalian intestine awaits further investigation. The present position is that the absence of CaBP in the intestine of rachitic animals of some species as claimed by some investigators[17,18] is open to doubt[19,20], although there is general acceptance that its synthesis is increased by the vitamin[17-21]. Resolution of this problem requires specific antisera to an intestinal CaBP preparation from a mammalian species in which vitamin D deficiency is readily produced.

R. Spencer
M. Charman
P. Wilson
E. Lawson

Dunn Nutritional Laboratory,
University of Cambridge and MRC,
Milton Road,
Cambridge CB4 1XJ, UK

1 Wasserman, R. H., Corradino, R. A., and Taylor, A. N., *J. biol. Chem.*, **243**, 3978–3986 (1968).
2 Wasserman, R. H., Taylor, A. N., and Fullmer, C. S., *Biochem. Soc. Spec. Publ.*, **3**, 55–74 (1974).
3 Emtage, J. S., Lawson, D. E. M., and Kodicek, E., *Nature*, **246**, 100–101 (1973).
4 Emtage, J. S., Lawson, D. E. M., and Kodicek, E., *Biochem. J.*, **144**, 339–346 (1974).
5 Ebel, J. G., Taylor, A. N., and Wasserman, R. H., *Am. J. clin. Nutr.*, **22**, 431–436 (1969).
6 Harmeyer, J., and DeLuca, H. F., *Archs Biochem. Biophys.*, **133**, 247–254 (1969).
7 Laurell, C. B., *Analyt. Biochem.*, **15**, 45–52 (1966).
8 Schachter, D., and Rosen, S. M., *Am. J. Physiol.*, **196**, 357–362 (1959).
9 Wasserman, R. H., and Taylor, A. N., *J. biol. Chem.*, **243**, 3987–3993 (1968).
10 Emtage, J. S., Lawson, D. E. M., and Kodicek, E., *Biochem. J.*, **140**, 239–247 (1974).
11 Hurwitz, S., Harrison, H. C., and Harrison, H. E., *J. Nutrit.*, **91**, 319–323 (1967).
12 Palmiter, R. D., Oka, T., and Schimke, R. T., *J. biol. Chem.*, **246**, 724–737 (1971).
13 Norman, A. W., *Vitams Horm.*, **32**, 325–384 (1974).
14 Corradino, R. A., *Science*, **179**, 402–404 (1973).
15 Wasserman, R. H., and Corradino, R. A., *Vitams Horm.*, **31**, 43–71 (1973).
16 Lippiello, L., and Wasserman, R. H., *J. Histochem. Cytochem.*, **23**, 111–116 (1975).
17 Freund, T., and Bronner, F., *Am. J. Physiol.*, **228**, 861–869 (1975).
18 Ooizumi, K., Moruichi, S., and Hosoya, N., *J. Vitamin.*, **16**, 228–234 (1970).
19 Harrison, J. E., Hitchman, A. J. W., and Tam, C. S., *Can. J. Physiol. Pharmac.*, **53**, 137–143 (1975).
20 Harrison, J. E., Hitchman, A. J. W., and Brown, R. G., *Can. J. Physiol. Pharmac.*, **53**, 144–149 (1975).
21 Drescher, N., and DeLuca, H. F., *Biochemistry*, **10**, 2302–2312 (1971).

Factors involved in initiation of haemoglobin synthesis can be phosphorylated *in vitro*

RABBIT reticulocytes, after incubation in a nutritional medium with radioactive inorganic phosphate, contain a significant amount of covalently bound phosphate. Approximately 65% of this phosphate is attached to ribosome-associated proteins which are released into the 0.5 M KCl wash fraction from ribosomes; the remainder is bound to ribosomal protein[1]. The high salt wash fraction contains various proteins required for the initiation and maintenance of protein synthesis. To analyse the effects of phosphorylation on the regulation of protein synthesis, purified initiation factors from rabbit reticulocytes have been examined for phosphate-acceptor activity through the use of cyclic AMP-regulated and cyclic nucleotide-independent protein kinases. Here we describe the phosphorylation of two factors previously shown to be involved in the initiation of haemoglobin synthesis. Both IF-MP, the Met-tRNA$_f$ binding factor, and IF-M2A, a ribosome-dependent GTPase activity, are modified by a cyclic nucleotide-independent protein kinase isolated from rabbit reticulocytes. Initiation factors IF-M1 (ref. 2), IF-M2B$_\alpha$ (ref. 3), IF-M2B$_\beta$ (ref. 3), IF-M3 (ref. 4) and IF-M4 (ref. 5) are not phosphorylated by this enzyme.

The cyclic nucleotide-independent protein kinase activities in the post-ribosomal supernatant fraction were isolated by chromatography on DEAE–cellulose (Fig. 1). Enzyme activity was monitored by incorporation of radioactive
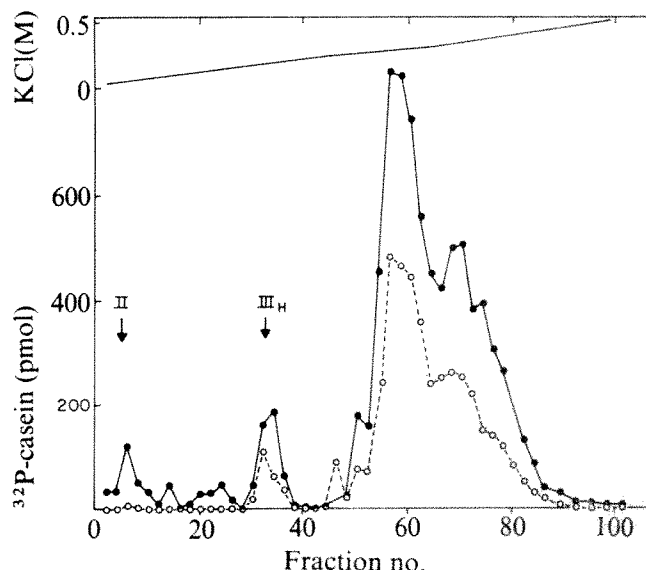
Fig. 1 Isolation of cyclic nucleotide-independent protein kinase activity from rabbit reticulocytes by chromatography on DEAE–cellulose. The post-ribosomal supernatant fraction was prepared and chromatographed on DEAE–cellulose as previously described[6] except that 600 mg of protein was added, and elution was with a 600-ml linear KCl gradient ranging from 0 to 0.5 M KCl. Protein kinase activity was assayed using γ-^{32}P-ATP (●) and γ-^{32}P-GTP (○) as previously described[6]. The activity referred to as III$_C$ in the text comprises fractions 54–79. Fractions 54–62, 67–72 and 74–79 were pooled individually, dialysed against 20 mM Tris-HCl, pH 7.5, 6 mM mercaptoethanol and 1 mM EDTA and stored at 4 °C.

phosphate from γ-^{32}P-ATP and γ-^{32}P-GTP into casein. The major protein kinase activity (III$_C$) elutes between 0.24 and 0.34 M KCl and is observed to contain several chromatographically different species. The protein kinase activity was pooled into three fractions as indicated in the legend to Fig. 1, and the enzyme activity in each of these fractions was analysed as a function of increasing concentrations of monovalent and divalent cations and for substrate specificity. No differences were observed among the three fractions (data not shown). The DEAE–cellulose column fractions were also analysed for cyclic AMP-regulated protein kinase activity using histone IIA (Sigma) as substrate in the presence of 1.4×10^{-6} M cyclic AMP. The regions where these two protein kinases eluted are indicated by arrows in Fig. 1. Further chromatography of the cyclic AMP-regulated protein kinases resolved these enzymes from cyclic nucleotide-independent activities[6].

Highly purified initiation factors from rabbit reticulocytes were incubated with the cyclic AMP-regulated and cyclic nucleotide-independent protein kinase activities using either radioactive ATP or GTP as the phosphate donor. Phosphate-acceptor activity of the initiation factors was demonstrated by electrophoresis of the incubation mixture on polyacrylamide gels containing sodium dodecyl sulphate (SDS) followed by autoradiography. IF-MP was phosphorylated as shown in Fig. 2. This initiation factor has been

Table 1 Effect of cyclic AMP and cyclic GMP on phosphorylation of IF-MP and IF-M2A

	IF-MP		IF-M2A	
	ATP	GTP	ATP	GTP
	^{32}P incorporated (pmol µg^{-1})			
Protein kinase (67–72)	2.7	1.9	11.6	8.0
+ cyclic AMP	2.4	1.9	13.6	9.7
+ cyclic GMP	2.5	2.1	12.1	10.7

The stained initiation factors obtained after polyacrylamide gel electrophoresis as described in Figs 2 and 3 were excised, dried and counted in a Beckman liquid scintillation counter with a toluene cocktail mixture.

purified to homogeneity[7], and has been shown to form a ternary complex with Met-tRNA$_f$ and GTP[8]. It is composed of three subunits which migrate with molecular weights of 57,000, 53,000 and 38,000 in polyacrylamide gels containing SDS[9]. The subunit with a molecular weight of 53,000 is phosphorylated by the cyclic nucleotide-independent protein kinase (III$_C$) using either ATP or GTP. With the reaction conditions used, 0.45 mol of phosphate were incorporated per mol of IF-MP, assuming a molecular weight of 180,000 (ref. 7). IF-MP is not modified by cyclic AMP-regulated protein kinases II or III$_H$.
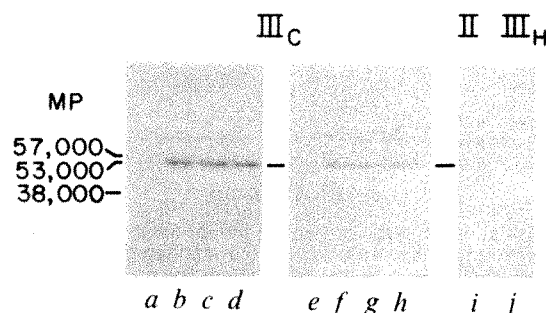
Fig. 2 Phosphorylation of IF–MP. IF-MP (1.4 μg) was phosphorylated in 0.045 ml of reaction mixture containing 50 mM Tris-HCl, pH 7.0; 10 mM MgCl$_2$, 0.14 mM γ-^{32}P-ATP (0.2 Ci mmol^{-1}) or γ-^{32}P-GTP (0.12 Ci mmol^{-1}); 1.4 × 10^{-6} M cyclic nucleotide (where indicated), and protein kinase. The reaction contained 7 enzyme units (EU) of III$_U$ (fractions 74–79), 25 EU of II, (further purified by chromatography on phosphocellulose[6]), or 20 EU of III$_H$ (purified further by chromatography on carboxymethyl Sephadex). A unit of protein kinase activity has been defined previously[6]. Incubation was for 20 min at 30 °C and terminated by the addition of SDS in sample buffer. Slab gel electrophoresis was performed using the gel system of Laemmli[21] as described by Ames[22] with the following modifications: the ratio of acrylamide to *bis*-acrylamide was 22.2 : 0.6; the electrode buffer contained twice the concentration of Tris-HCl and glycine; and the resolving gel and the stacking gel acrylamide concentrations were 10% and 5% respectively. Electrophoresis was at room temperature for 0.5 h at 100 V followed by 3.5 h at 150 V. The gels were stained in Coomassie brilliant blue, destained and autoradiographed for 3–4 h against Kodak No-Screen medical X-ray film. Samples (*a–d*) contained ATP and III$_C$. *a*, IF-MP omitted; *b*, no additions; *c*, plus cyclic AMP; *d*, plus cyclic GMP. Samples (*e–h*) contained GTP and III$_C$. *e*, IF-MP omitted; *f*, no additions; *g*, plus cyclic AMP; *h*, plus cyclic GMP. *i*, Contained ATP, cyclic AMP, and II; *j*, contained ATP, cyclic AMP and III$_H$.

Fig. 3 Phosphorylation of IF-M2A. The procedures were described for Fig. 2 except that 1.5 μg of homogeneous IF-M2A was reacted with 3 EU of protein kinase III$_C$ (fractions 67–72) in a final volume of 0.025 ml for 10 min. The stacking gel contained 2.5% acrylamide and the voltage for the first 0.5 h of the run was 75 V. For identification see Fig. 2.

IF-M2A which catalyses a ribosome-dependent GTPase activity was also modified by the cyclic nucleotide-independent protein kinase activity that phosphorylates IF-MP (Fig. 3). This initiation factor has been observed to stimulate protein synthesis using either a natural or artificial template and has been purified to homogeneity[10]. As observed in the autoradiogram, either ATP or GTP can be the phosphate donor. With optimal incubation conditions, approximately two phosphoryl moieties can be incorporated per polypeptide of molecular weight 125,000, and both serine and threonine residues are modified (data not shown). IF-M2A is not modified by either of the cyclic AMP-regulated protein kinases.

The effects of cyclic AMP and cyclic GMP on the phosphorylation of both initiation factors are shown qualitatively by autoradiography in Figs 2 and 3. To quantify the effects of these nucleotides on the phosphorylation of the initiation factors, the radioactive proteins were excised from the gel, dried and counted in a liquid scintillation counter. Table 1 shows that cyclic AMP and

cyclic GMP neither stimulate nor inhibit the phosphorylation of IF-MP and IF-M2A at 10^{-6} M. In the conditions of the reaction, ATP is slightly more effective as a phosphate donor than GTP. Both the ATP and GTP were examined by thin-layer chromatography on PEI cellulose and found to be chromatographically pure with regard to contaminating radioactive nucleotides. Thus, both nucleotide triphosphates are authentic phosphate donor compounds in the reaction. Both initiation factors are modified by the three chromatographically different fractions of protein kinase which comprise III$_C$. These three fractions show little or no difference in the extent to which they phosphorylate the initiation factors.

Our results provide the first evidence of a cyclic nucleotide-independent phosphorylation of initiation factors which are involved in the complex process of protein synthesis. The phosphorylation of a single 40S ribosomal protein in rabbit reticulocytes has been shown to be controlled by cyclic AMP[1,11]. Phosphorylation of the same protein in rat liver increases considerably when the liver is forced to regenerate[12] or in conditions of experimental diabetes[13].

Evidence suggests that the role of haemin in the regulation of protein synthesis is mediated by phosphorylation. In the absence of added haemin, inhibition of protein synthesis in a reticulocyte lysate occurs at chain initiation with a concomitant decrease in binding of Met-tRNA$_f$ to 40S ribosomal subunits[14,15]. This inhibition is potentiated by the addition of ATP[16,17]. Inhibition produced by haemin deprivation is prevented in the reticulocyte lysate by high concentrations of GTP[16,17], cyclic AMP[17,18] and various purines[18], and by the addition of IF-MP[19]. A protein kinase activity has been identified with the purified inhibitor fraction[20]. Since the protein kinase activity described here makes use of both ATP and GTP as substrate, it is questionable whether this enzyme has the requisite properties for the translational inhibitor. But we have recently isolated another cyclic nucleotide-independent protein kinase activity which initially copurifies with cyclic AMP-regulated protein kinase III$_H$. This cyclic nucleotide-independent activity modifies the small subunit of IF-MP and may be the haem-controlled repressor (S.M.T., J.A.T., B.S. and W.C.M., in preparation).

Thus there are at least two possible mechanisms for the control of translation by phosphorylation—cyclic AMP-regulated modification of the 40S ribosomal subunit and cyclic nucleotide-independent phosphorylation of the initiation factors. We are analysing the differential phosphorylation of the two factors and the 40S ribosomal subunit in order to understand the molecular basis for the control of protein synthesis by phosphorylation.

JOLINDA A. TRAUGH
STANLEY M. TAHARA
SANDRA B. SHARP

Department of Biochemistry,
University of California,
Riverside, California 92502

BRIAN SAFER
WILLIAM C. MERRICK

Molecular Hematology Branch,
National Heart and Lung Institute,
National Institutes of Health,
Bethesda, Maryland 20014

1 Traugh, J. A., and Porter, G. G., *Biochemistry*, 15, 610–616 (1976).
2 Merrick, W. C., and Anderson, W. F., *J. biol. Chem.*, 250, 1197–1206 (1975).
3 Kemper, W. M., Berry, K. W., and Merrick, W. C., *J. biol. Chem.*, 251, (in the press).
4 Prichard, P. M., and Anderson, W. F., *Meth. Enzymol.*, 30, 136–141 (1974).
5 Safer, B., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, (in the press).
6 Traugh, J. A., and Traut, R. R., *J. biol. Chem.*, 249, 1207–1212 (1974).
7 Safer, B., Anderson, W. F., and Merrick, W. C., *J. biol. Chem.*, 250, 9067–9075 (1975).
8 Safer, B., Adams, S. L., Anderson, W. F., and Merrick, W. C., *J. biol. Chem.*, 250, 9083–9089 (1975).
9 Tahara, S. M., Traugh, J. A., Safer, B., and Merrick, W. C., *Fedn Proc.*, 35, 1515 (1976).
10 Merrick, W. C., Kemper, W. M., and Anderson, W. F., *J. biol. Chem.*, 250, 5556–5562 (1975).
11 Cawthon, M. L., Bitte, L. F., Krystosek, A., and Kabat, D., *J. biol. Chem.*, 249, 275–278 (1974).
12 Gressner, A. M., and Wool, I. G., *J. biol. Chem.*, 249, 6917–6925 (1974).
13 Gressner, A. M., and Wool, I. G., *Nature*, 259, 148–150 (1976).
14 Balkow, K., Mizuno, S., Fisher, J. M., and Rabinovitz, M., *Biochim. biophys. Acta*, 324, 397–409 (1973).
15 Legon, S., Jackson, R. J., and Hunt, T., *Nature new Biol.*, 241, 150–152 (1973).
16 Balkow, K., Hunt, T., and Jackson, R. J., *Biochem. biophys. Res. Commun.*, 67, 366–375 (1975).
17 Ernst, V., Levin, D. H., Ranu, R. S., and London, I. M., *Proc. natn. Acad. Sci. U.S.A.*, 73, 1112–1116 (1976).
18 Legon, S., Brayley, A., Hunt, T., and Jackson, R. J., *Biochem. biophys. Res. Commun.*, 56, 745–752 (1974).
19 Kaempfer, R., *Biochem. biophys. Res. Commun.*, 61 591–597 (1974).
20 Levin, D. H., Ranu, R. S., Ernst, V., Fifer, M. A., and London, I. M., *Proc. natn. Acad. Sci. U.S.A.*, 72, 4849–4853 (1975).
21 Laemmli, U. K., *Nature*, 227, 680–685 (1970).
22 Ames, G. F. L., *J. biol. Chem.*, 249, 634–644 (1974).

Evidence for involvement of nuclear envelope nucleoside triphosphatase in nucleocytoplasmic translocation of ribonucleoprotein

A NUCLEOSIDE triphosphatase activity showing broad substrate specificity has been found in nuclear envelope isolated from rat and pig liver, *Tetrahymena* macronucleus, and cultured SV-3T3 fibroblasts using the method of Harris and Milne[1], and has been extensively characterised in this laboratory. There is evidence that in purified nuclear envelopes there is only one enzyme which catalyses nucleoside triphosphate hydrolysis: the activities observed with different substrates and metal ions are not additive, and they show a single pH optimum, the value of which varies with the source of the envelopes[2]. This enzyme seems to be localised in the nuclear pore complexes[3]. Several reports[4–13] have indicated that ATP markedly stimulates the release of mRNA from isolated nuclei, but no evidence has been presented to suggest that this stimulation necessarily requires hydrolysis of the ATP. In this paper, however, we present evidence that the nucleoside triphosphatase is an essential component of the mechanism of nucleocytoplasmic translocation of ribonucleoprotein.

SV-3T3 cells were grown in Dulbecco's modified Eagle's medium (Flow Laboratories) and labelled for 6 h with 10 μCi $^{14}C_6$-orotic acid (61 mCi mmol^{-1}, Amersham). The cells were homogenised in 10 mM Tris-HCl, pH 7.4, containing 1 mM MgCl$_2$, and centrifuged through 2.2 M sucrose in the same Tris–MgCl$_2$ buffer at 100,000g for 1 h at 0 °C. The pellet of nuclei was substantially free of 5′ nucleotidase
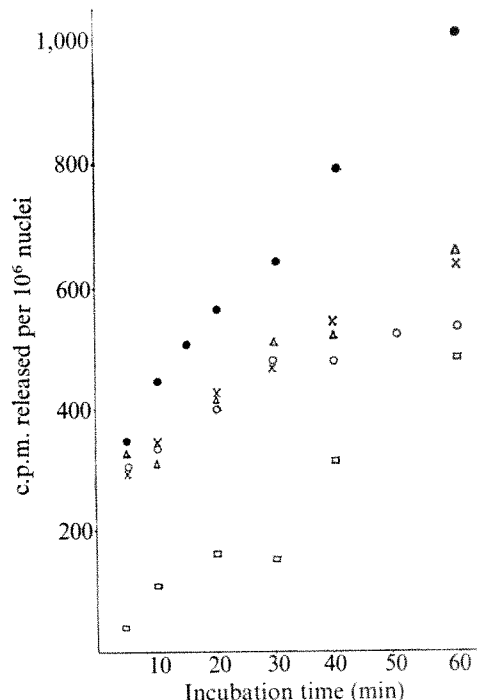
Fig. 1 Time course of RNA release from isolated nuclei. All experiments were carried out at 35 °C in the following medium[13]: 50 mM Tris-HCl, pH 7.6, 25 mM KCl, 2.5 mM MgCl$_2$, 0.3 mM MnCl$_2$, 0.5 mM CaCl$_2$, 5.0 mM NaCl, 2.5 mM K$_2$HPO$_4$, 5.0 mM spermidine and 2.0 mM dithioerythrytol. ○, Control; ●, release on addition of 2.0 mM ATP; ×, release on addition of 2.0 mM EDTA; △, release on addition of 2.0 mM adenylyl-β,γ-methylene diphosphonate (AMP.PCP); □, difference between ATP-stimulated and control values. After 30 min incubation, $74 \pm 4\%$ of the counts released in the control and $90 \pm 5\%$ of the additional counts released in the presence of ATP were precipitable with 5% TCA. There was no significant difference between the release in the presence of EDTA and that in the presence of AMP.PCP; indicating that, contrary to the finding of Raskas[7], the effect of AMP.PCP is solely attributable to its chelating action. The effect of ATP, however, is clearly not wholly attributable to chelation.

(EC 3.1.3.5), Na/K-stimulated Mg-ATPase (EC 3.6.1.4) and succinate dehydrogenase (EC 1.3.99.1), and no organelle contamination was visible under phase contrast (B.McC., P.S.A. and J. F. Lamb, unpublished). Digestion of the nuclei in 0.1 M NaOH and counting in a Packard Tri-Carb scintillation spectrophotometer using NE 250 scintillant (NEM) showed ^{14}C incorporation of $5,000 \pm 1,000$ c.p.m. per 10^6 nuclei. In the presence of 50 μg ml^{-1} proflavine, orotic

Table 1 Specificity of the nucleoside triphosphatase and of nucleoside-triphosphate-stimulated RNA release from isolated nuclei for various divalent cation–nucleoside triphosphate complexes

Substrate	a	b
Mg.ATP^{2-}	16.7 ± 1.1	438 ± 21
Mg.GTP^{2-}	18.0 ± 1.1	450 ± 22
Mg.UTP^{-}	7.9 ± 0.5	219 ± 15
Mg.CTP^{2-}	7.2 ± 0.5	153 ± 13
Ca.ATP^{2-}	13.4 ± 1.1	350 ± 19
Mn.ATP^{2-}	11.5 ± 1.1	306 ± 18
Zn.ATP^{2-}	1.3 ± 0.4	22 ± 10
Be.ATP^{2-}	Zero	Trace

a, Nucleoside triphosphatase activity (μmol inorganic phosphate released per h per mg protein). *b*, Stimulated RNA release (c.p.m. per h per 10^6 nuclei)—that is, the difference between counts released in the presence and in the absence of the nucleoside triphosphate. For the release experiments, nuclei were incubated at 35 °C in 0.88 M sucrose containing either 5 mM MgCl$_2$, CaCl$_2$, MnCl$_2$, ZnCl$_2$ or BeCl$_2$, and the 5% TCA precipitate of the supernatant was counted as described in the text[4]. The results in column *b* are in good agreement with those of Ishikawa *et al.*[4].
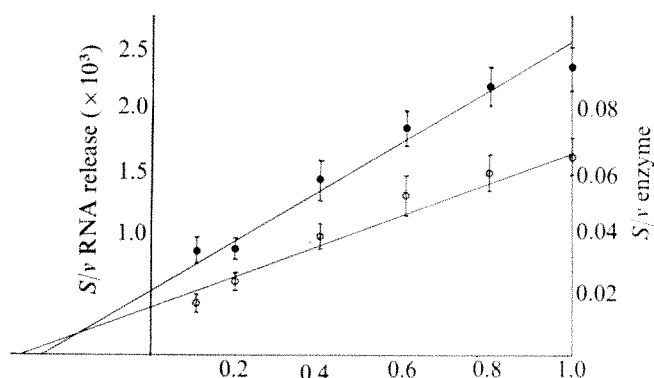
Fig. 2 Eadie plots comparing dependence of rates of nucleoside triphosphatase reaction (○) and ATP-stimulated RNA release from nuclei (●) on ATP concentration. For the left-hand ordinate (S/v RNA release), the v term has been taken as the difference in label released (c.p.m. per h per 10^6 nuclei) in the presence and absence of ATP at each concentration. For the right-hand ordinate (S/v enzyme), the v term is the nucleoside triphosphatase activity (μmol phosphate released per h per mg envelope protein). Enzyme assays were carried out as described in the text, the free Mg^{2+} concentration being maintained at 0.2 mM throughout[2,18]. RNA release experiments were carried out essentially as described in Fig. 1; but to avoid spurious results due to the chelating action of ATP, EDTA was added to make the total ATP+EDTA concentration 1.0 mM throughout. K_m values obtained are 0.26 ± 0.05 mM for the nucleoside triphosphatase, and 0.30 ± 0.09 mM for the ATP-stimulated RNA release. The agreement between the behaviour of the two processes is again close.

acid incorporation was reduced to 850 c.p.m. per 10^6 nuclei, indicating that most of the label was taken up into the nuclear RNA[14].

The media described by Ishikawa *et al.*[4] and by McNamara *et al.*[13], which maintain the nuclei in a stable and essentially disaggregated state, were used in the RNA release experiments; the yeast RNA, ATP-generating system and cytoplasmic protein[13] were omitted. Volumes (0.75 ml) of a suspension containing approximately 5×10^6 nuclei ml^{-1}, with and without ATP (routinely 2.0 mM) and other reagents, were incubated at 35 °C for various times. The suspension was centrifuged at $3,000g$ for 5 min to remove nuclei, and 0.4-ml aliquots of the supernatant counted as before.

The nucleoside triphosphatase activity proved resistant to all attempts at solubilisation. Enzyme assays were therefore carried out by incubating whole nuclear envelopes in 25 mM triethanolamine–HCl buffer, pH 7.7, with ATP and $MgCl_2$

Table 2 Effects of various inhibitors on the nuclear envelope nucleoside triphosphatase and on ATP-stimulated RNA release from isolated nuclei

Inhibitor	a	b
Ouabain 2.0 mM	100 ± 8	100 ± 9
Gramicidin 50 μg ml^{-1}	100 ± 6	100 ± 3
Cycloheximide 50 μg mL^{-1}	100 ± 4	100 ± 7
Oligomycin 5 μg ml^{-1}	85 ± 5	82 ± 8
50 μg ml^{-1}	60 ± 10	52 ± 6
Quercetin 5 μg ml^{-1}	50 ± 6	35 ± 5
15 μg ml^{-1}	Trace	Trace
AMP.PCP 2.0 mM	100 ± 7	100 ± 6
ATP γ S 1.0 mM	50 ± 10	38 ± 6
2.0 mM	16 ± 8	Trace

a, Nucleoside triphosphatase activity; b, rate of RNA release. All values are expressed as percentage of control ±1 s.e.m. ATP γ S, Adenosine-5'-O-(3)-thiotriphosphate. ATP analogues were obtained from Boehringer; other inhibitors from Sigma.

at 35 °C for 30 min, and determining inorganic phosphate release[15]. In experiments in which the effects of inhibitors were investigated, the nuclei or envelopes were preincubated with the inhibitor for 5 min, and solvent blanks were used as controls.

Substrate specificity, kinetic behaviour and inhibition properties of the ATP-stimulated RNA release were investigated and compared with the corresponding properties of the nucleoside triphosphatase, to establish whether there was a causal connection between the two processes. The effect of ATP on RNA release from the nuclei is shown in Fig. 1: 25–30% of the label released in the absence of exogenous ATP is attributable to low molecular weight (TCA-soluble) material; the remainder presumably represents leakage of RNA from the nuclei, although the possibility that traces of endogenous ATP may contribute to the background cannot be ruled out. Preliminary isotachophoretic measurements indicate not more than 20 nmol ATP per 10^6 freshly prepared nuclei.

Table 1 shows that replacement of other nucleoside triphosphates for ATP and of other divalent cations for Mg^{2+} still resulted in stimulation of RNA release from the nuclei. There is a striking correlation between this stimulated release and the rate of hydrolysis catalysed by the nucleoside triphosphatase for all divalent cation–nucleoside triphosphate complexes investigated. The nuclear envelope contains an active glucose-6-phosphatase (EC 3.1.3.9) (ref. 3), but replacement of ATP by 2.0 mM glucose-6-phosphate was not attended by significant stimulation of the RNA release.

Table 2 compares the effects of inhibitors on the nucleoside triphosphatase and on the ATP-stimulated release of RNA. Again, the correlation is striking, particularly so because the inhibitor concentrations for half-maximal inhibition were similar for both processes. The dependence of the rate of ATP-stimulated RNA release on ATP concentration approximated to Michaelis–Menten kinetics, with a K_m' value $(2.8\times10^{-4}$ M) close to that for the nucleoside triphosphatase (Fig. 2). For all properties investigated, therefore, there was a close correspondence between the behaviour of the two processes, indicating a likely causal connection.

There is strong evidence that the RNA released as a ribonucleoprotein complex into the supernatant in the conditions described in this paper is predominantly mRNA. Although the possibility of some contamination with heterogeneous nuclear RNA cannot be ruled out, normal nuclear restriction of RNA has been clearly demonstrated using the medium described by McNamara *et al.*[16] and Ishikawa *et al.* note that the RNA released in the conditions they use in the presence of ATP is largely incorporated into polysomes[17]. If these *in vitro* conditions, which have been used in the present investigation, can be accepted as a reasonable facsimile of conditions *in vivo*, then the results we have described here strongly suggest that the nuclear envelope nucleoside triphosphatase constitutes an essential part of the mechanism of nucleocytoplasmic ribonucleoprotein translocation. Detailed kinetic studies of the enzyme have been carried out in this laboratory[18], with a view to illuminating the molecular mechanism of this translocation process.

P. S. Agutter
H. J. McArdle
B. McCaldin

Department of Physiology,
Bute Medical Buildings,
University of St Andrews,
St Andrews, Fife, UK

1 Harris, J. R., and Milne, J. F., *Biochem. Soc. Trans.*, **2**, 1251-1253 (1974).
2 Agutter, P. S., *et al.*, *Biochim. biophys. Acta* (in the press).
3 Franke, W. W., *Int. Rev. Cytol. suppl.*, **4**, 71-236 (1974).
4 Ishikawa, K., Kuroda, C., and Ogata, K., *Biochim. biophys. Acta*, **179**, 316-331 (1969).
5 Ishikawa, K., Kuroda, C., Ueki, M., and Ogata, K., *Biochim. biophys. Acta*, **213**, 495-504 (1970).
6 Ishikawa, K., Kuroda, C., and Ogata, K., *Biochim. biophys. Acta*, **213**, 505-512 (1970).
7 Raskas, H. J., *Nature new Biol.*, **233**, 134-136 (1971).
8 Raskas, H. J., and Bhaduri, S., *Biochemistry*, **12**, 920-925 (1973).
9 Schumm, D. E., and Webb, T. E., *Biochim. biophys. Res. Commun.*, **48**, 1259-1265 (1972).
10 Yu Ling, C., Racevskis, J., and Webb, T. E., *Cancer Res.*, **32**, 2314-2321 (1972).
11 Schumm, D. E., McNamara, D. S., and Webb, T. E., *Nature new Biol.*, **245**, 201-203 (1973).
12 Schumm, D. E., Morris, H., and Webb, T. E., *Cancer Res.*, **33**, 1821-1828 (1973).
13 McNamara, D. S., Racevskis, J., Schumm, D. E., and Webb, T. E., *Biochem. J.*, **147**, 193-197 (1975).
14 Stenram, U., and Willem, R., *Expl Cell Res.*, **50**, 505-514 (1968).
15 Wohler, B. E., and Wollenberger, A., *Biochim. Z.*, **329**, 508-520 (1958).
16 Schumm, D. E., and Webb, T. E., *Biochem. biophys. Res. Commun.*, **58**, 354-360 (1974).
17 Ishikawa, K., Ueki, M., Nagai, K., and Ogata, K., *Biochim. biophys. Acta*, **259**, 138-154 (1972).
18 Agutter, P. S., Cockrill, J. B., Harris, J. R., Vincent, C. A., and Wolfe, M. A., *Biochim. biophys. Acta* (in the press).

Unexpected occurrence of an aminoacylated nucleoside in mammalian tRNATyr

MANY tRNAs contain a hypermodified nucleoside at the 3' side of their anticodons[1]. The modifications at this position are essential for tRNA function in ribosome binding and amino acid transfer[2-4], and cause inability to base pair according to the Watson–Crick scheme[5]. A striking correlation between the codon response of the tRNA and the nature of this hypermodified nucleoside has been observed: all known sequences of tRNAs (*Escherichia coli*), with codons beginning with uracil, contain N^6-(Δ^2-isopentenyl)-2-methylthio-adenosine at the 3' side of the anticodon[6-8], whereas tRNAs with codons starting with adenine, contain an aminoacylated nucleoside, such as N-[N-(9-β-D-ribofuranosyl-purin-6-yl)carbamoyl] threonine (t^6A) or the methyl derivative (mt^6A) at this position[9,10].

The same distribution of these isopentenylated and aminoacylated nucleosides has also been observed in yeast[11-13] and mammals[14,15]. This striking regularity led to the widely accepted postulate that N^6-(Δ^2-isopentenyl-adenosine) (i^6A) or the methylthiolated derivative might have a specific role in the recognition of codons starting with uridine[1], whereas t^6A is thought to prevent pairing between uridine in the third position of the anticodon and guanine in the first position of the codon[16].

The Tyr-tRNAs read the UAX-codons and contain in *E. coli*[17,18], *Torulopsis utilis*[19], and yeast[11] i^6A next to the anticodon, as expected. Tyr-tRNA from rat liver, however, chromatographs very differently from the i^6A-containing species on partition chromatography columns[20], and elutes at about the same position as tRNA$_3^{Ser}$ which contains the hydrophilic nucleoside mt^6A (ref. 15). This chromatographic behaviour

of tRNA^{Tyr} (rat liver) indicates the presence of a polar nucleoside in the anticodon loop of this molecule. As mentioned, t^6A was thought to be present exclusively in tRNAs coding for AXX codons, so we did not expect it in tRNA^{Tyr}, which codes for UAX. Here we report the unexpected occurrence of t^6A in tRNA^{Tyr} from rat liver.

tRNA^{Tyr} was purified from crude rat liver tRNA (ref. 21) using two different fractionation procedures. The first involved fractionation on BD–cellulose column chromatography[22] at pH 4.5. The tRNA^{Tyr}-enriched fractions were then separated on a DEAE–Sephadex A 50 column[23] at pH 7.3, further purified on a BD–cellulose column in 7 M urea[24] at pH 4.5, and finally separated on a reversed phase column (system 5)[25]. The second purification method consisted of fractionation on a BD–cellulose column in 7 M urea at pH 4.5, then the tRNA^{Tyr}-enriched fractions were phenoxyacylated[26] and rechromatographed using the same chromatography system. The latter method takes advantage of the fact that a modified uridine in tRNA^{Tyr} reacts specifically with the phenoxyacylating reagent leading to a more lipophilic elution position of the derivatised tRNA, so that tRNA^{Tyr} can be separated rapidly from non-derivatised material[27]. From the elution profile of other tRNAs one can conclude that the structures of hypermodified nucleosides in the anticodon loop are not affected by this method.

After enzymatic hydrolysis to nucleosides, pure fractions of tRNA^{Tyr} (rat liver) were analysed by two-dimensional thin-layer chromatography on cellulose-coated aluminium foils, as described previously[28]. The nucleosides were identified by their characteristic chromatographic behaviour and by their ultraviolet absorption spectra at pH 1, 6 and 12, as well as by their characteristic fluorescence and phosphorescence at 77 K (ref. 29). The two-dimensional chromatogram of a hydrolysate of rat liver tRNA^{Tyr} is shown in Fig. 1. With the exception of the chemically modified uridine derivative in case of the phenoxyacylated tRNA, identical patterns were obtained by analysis of fractions from the two different purification methods. The following nucleosides were found: adenosine, 13-15; cytidine, 19-20; guanosine, 20-22; uridine, 10-11; pseudouridine, 3-4; ribothymidine, 1; dihydrouridine, 3-4; 1-methyladenosine, 1; 5-methylcytidine, 1; 1-methylguanosine, 1; N^2-methylguanosine, 1; N^2,N^2-dimethylguanosine, 1; 7-methylguanosine, 1; and an unknown uridine derivative (U*) which reacts with the phenoxyacylating reagent. Neither i^6A, nor one of its derivatives, was found in the hydrolysate. Instead, one residue of the nucleoside t^6A, which is found exclusively at the 3' position adjacent to the anticodon, was detected. Since t^6A was thought to be present only in tRNAs coding for AXX codons and therefore should not occur in the tyrosine-specific tRNA, we proved the identity of t^6A by the following extensive criteria. (1) The ultraviolet absorption spectra of the compound at different pH showed the same characteristics as those of the synthetic t^6A (Fig. 2). (2) After alkaline hydrolysis of t^6A from tRNA^{Tyr} with KOH (0.3 M, 16 h, 37 °C)
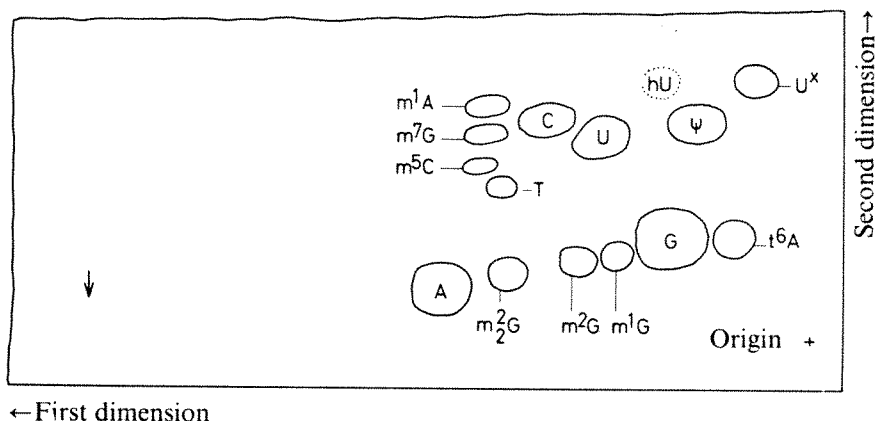
Fig. 1 A two-dimensional thin-layer chromatogram of the digest of tRNA^{Tyr}. 2–3 A_{260} units of the desalted tRNA^{Tyr} were degradated enzymatically overnight at 37 °C by 25 μg pancreatic RNase (Boehringer GmbH, Mannheim), 30 μg snake venom phosphodiesterase (Worthington) and 25 μg alkaline phosphatase (Worthington, Code BAPF) in a siliconised test tube containing 0.02 M ammonium formate (pH 7.6) and 5×10^{-4} M MgCl$_2$ in 0.2 ml (ref. 28). The digest was subsequently treated for 2 h at 37 °C with 5 μg ribonuclease T$_1$ (Calbiochem) and then applied on a cellulose-coated aluminium foil, as described previously[28]. The solvent systems used (see Table 1) were solvent C in the first dimension and solvent D in the second. The arrow indicates the established position of isopentenyladenosine which is absent in this hydrolysate.

←First dimension

Second dimension →
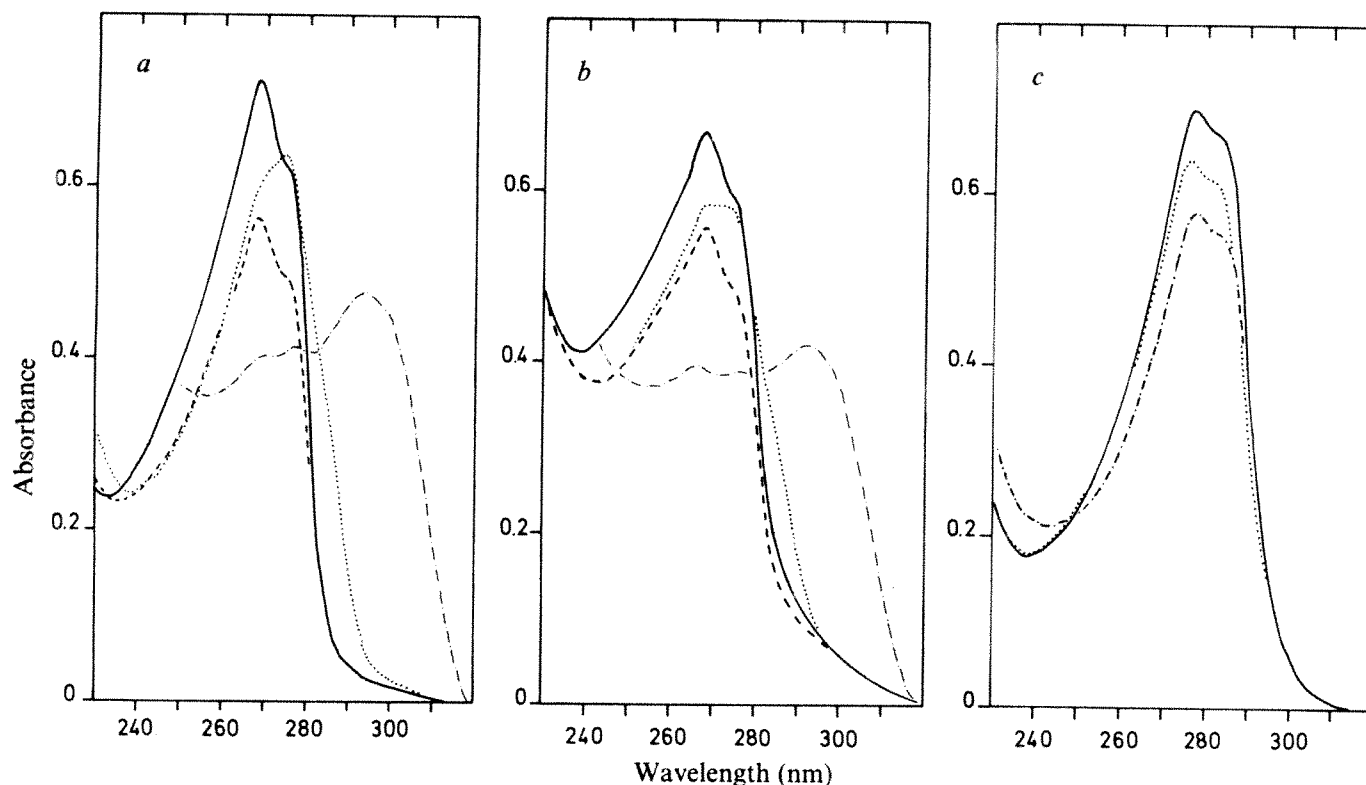
Fig. 2 Ultraviolet absorption spectra of aminoacylated nucleosides. *a*, Synthetic t⁶A; *b*, t⁶A from rat liver tyrosine tRNA; *c*, synthetic mt⁶A.
···· pH 1; —— pH 6; — — — pH 12; —·—·— pH 13.

and removal of the potassium with perchloric acid, the ultraviolet absorption spectra of the hydrolysate was identical to that of adenosine. In addition, the degradation product also showed mobility identical to adenosine on three different thin-layer chromatographic systems. (3) Automatic amino acid analysis after degradation of the compound in $Ba(OH)_2$ (0.1 M, 20 min in a boiling water bath and 16 h at 37 °C, followed by removal of Ba^{2+} with sulphuric acid) yielded equal amounts of threonine and serine. Control experiments using synthetic t⁶A established that in these hydrolysis, considerable amounts of threonine were degraded to serine. (4) The chromatographic properties of t⁶A from tRNA^{Tyr} in four different solvent systems were identical to those of synthetic t⁶A and different from either its glycine analogue g⁶A or mt⁶A (Table 1). (5) Chromatography on DEAE-cellulose at pH 8–8.8 showed that the compound had a negative charge like t⁶A and differed therefore from another aminoacylated derivative which had been isolated from tRNA$_2$^{Met} (*E. coli*)[1] and which would be uncharged at this pH.

These findings clearly prove the presence of the aminoacylated nucleoside t⁶A in tRNA^{Tyr} (rat liver). Analysis of

Table 1 Chromatographic mobilities of aminoacylated nucleosides compared with t⁶A from tRNA^{Tyr} (rat liver) in four different solvent systems

Nucleoside	R_F values ($\times$ 100) in solvent			
	A	B	C	D
t⁶A from tRNA^{Tyr} (rat liver)	49	53	10	30
t⁶A (synthetic)	49	53	10	30
mt⁶A	58	59	13	27
g⁶A	38	48	5	10
adenosine	54	74	42	11

The following solvent systems (cited in ref. 28) were used: A, *n*-propanol–concentrated ammonium hydroxide–water (90:45:15, by volume); B, isobutyric acid–ammonium hydroxide (0.5 M)–EDTA (0.1 M) (100:60:1.6, by volume); C, 1-butanol–isobutyric acid–water–ammonium hydroxide (25%) (150:75:50:5, by volume) D, saturated ammonium sulphate–sodium acetate (0.1 M) (pH 6)–isopropanol (79:19:2, by volume).

tRNA^{Tyr} from the silk worm (*Bombyx mori*), showed that in tRNA^{Tyr} from this organism, i⁶A is also replaced by t⁶A, indicating that t⁶A in tRNA^{Tyr} seems to be a common feature of higher organisms—thus contrasting strongly the situation in *E. coli*[6–8] and yeast[11–12].

It is particularly interesting that the nature of the base adjacent to the anticodon of tRNAs coding for UXX codons is very different in *E. coli* and in mammals. In *E. coli* this position is always occupied by N^6-(Δ^2-isopentenyl)-2-methylthio-adenosine, whereas in higher organisms there were detected (in addition to isopentenyl-adenosine) wybutosine (Y) in tRNA^{Phe} (yeast)[30], 1-methyl-guanosine in tRNA$_3$^{Leu} (yeast)[31], adenosine in tRNA^{Trp} (yeast)[32] and derivatives of wybutosine in tRNA^{Phe} of wheat germ[33] and mammals[34–37]. This may indicate that the problem of recognition of these tRNAs by the ribosome might be solved differently in prokaryotes and higher organisms.

We thank Drs G. B. Chheda and M. P. Schweizer for synthetic aminoacylated nucleosides (t⁶A, mt⁶A, g⁶A), Dr J.-P. Garel for the sample of pure tRNA^{Tyr} from *Bombyx mori*, and Drs J. Gordon, G. A. Howard and J. P. Jost for critical reading of the manuscript.

R. BRAMBILLA
H. ROGG
M. STAEHELIN

Friedrich Miescher-Institut,
PO Box 273
CH-4002 Basel, Switzerland

1 Nishimura, S., in *Progress in Nucleic Acid Research and Molecular Biology*, 12 (edit. by Davidson, J. N., and Cohn, W. E.), 50–85 (Academic, New York, 1972).
2 Gefter, M. L., and Russell, R. L., *J. molec. Biol.*, 39, 145–157, (1969).
3 Fittler, F., and Hall, R. H., *Biochem. biophys. Res. Commun.*, 25, 441–446 (1966).
4 Thiebe, R., and Zachau, H. G., *Eur. J. Biochem.*, 5, 546–555 (1968).
5 Parthasarathy, R., Ohrt, J. M., and Chheda, G. B., *Biochem. biophys. Res. Commun.*, 60, 211–218 (1974).
6 Harada, F., *et al.*, *Biochem. biophys. Res. Commun.*, 33, 299–306 (1968).
7 Nishimura, S., Yamada, Y., and Ishikura, H., *Biochim. biophys. Acta*, 179, 517–520 (1969).
8 Yamada, Y., Nishimura, S., and Ishikura, H., *Biochim. biophys. Acta*, 247, 170–174 (1971).

9 Ishikura, H., Yamada, Y., Murao, K., Saneyoshi, M., and Nishimura, S., *Biochem. biophys. Res. Commun.*, 37, 990–995 (1969)
10 Kimura-Harada, F., Harada, F., and Nishimura, S., *FEBS Lett.*, 21, 71–74 (1972).
11 Madison, J. T., and Kung, H. K., *J. biol. Chem.*, 242, 1324–1330 (1967).
12 Smith, C. J., Teh, H. S., Leu, A. N., and D'Obrenan, P., *J. biol. Chem.*, 248, 4475–4485 (1973).
13 Kuntzel, B., Weissenbach, J., and Dirheimer, G., *FEBS Lett.*, 25, 189–191 (1972).
14 Staehelin, M., Rogg, H., Baguley, B. C., Ginsberg, T., and Wehrli, W., *Nature*, 219, 1363–1365 (1968).
15 Rogg, H., Müller, P., and Staehelin, M., *Eur. J. Biochem.*, 53, 115–127 (1975).
16 Jukes, T. H., *Nature*, 246, 22–26 (1973).
17 Goodman, H. M., Abelson, J., Landy, A., Brenner, S., and Smith, J. D., *Nature*, 217, 1019–1024 (1968).
18 Goodman, H. M., Abelson, J., Landy, A., Zadrazil, S., and Smith, J. D., *Eur. J. Biochem.*, 13, 461–483 (1970).
19 Hashimoto, S., Miyazaki, M., Takemura, S., *J. Biochem.*, *Tokyo*, 65, 659–661 (1969).
20 Wehrli, W., and Staehelin, M., *Biochemistry*, 10, 1878–1884 (1971).
21 Rogg, H., Wehrli, W., and Staehelin, M., *Biochim. biophys. Acta*, 195, 13–15 (1969).
22 Gillam, I., et al., *Biochemistry*, 6, 3043–3056 (1967).
23 Cherayil, J. D., and Bock, R. M., *Biochemistry*, 4, 1174–1183 (1965).
24 Fittler, F., Kruppa, J., and Zachau, H. G., *Biochim. biophys. Acta*, 277, 513–522 (1972).
25 Pearson, R. L., Weiss, J. F., and Kelmers, A. D., *Biochim. biophys. Acta*, 228, 770–774 (1971).
26 Gillam, I., Blew, D., Warrington, R. C., von Tigerstrom, M., and Tener, G. M., *Biochemistry*, 7, 3459–3468 (1968).
27 Friedman, S., *Biochemistry*, 11, 3435–3443 (1972).
28 Rogg, H., Brambilla, R., Keith, G., and Staehelin, M., *Nucleic Acids Res.*, 3, 285–295 (1976).
29 Rogg, H., Wehrli, W., and Staehelin, M., in *Methods in Enzymology*, 20 (edit. by Moldave, K., and Grossman, L.), 118–125 (Academic, New York, 1971).
30 RajBhandary, U. L., et al., *Proc. natn. Acad. Sci. U.S.A.*, 57, 751–758 (1967).
31 Chang, S. H., Kuo, S., Hawkins, E., and Miller, N. R., *Biochem. biophys. Res. Commun.*, 51, 951–955 (1973).
32 Keith, G., Roy, A., Ebel, J.-P., and Dirheimer, G., *FEBS Lett.*, 17, 306–308 (1971).
33 Dudock, B. S., Katz, G., Taylor, E. K., and Holley, R. W., *Proc. natn. Acad. Sci. U.S.A.*, 62, 941–945 (1969).
34 Keith, G., et al., *FEBS Lett.*, 31, 345–347 (1973).
35 Keith, G., Ebel, J.-P., and Dirheimer, G., *FEBS Lett.*, 48, 50–52 (1974).
36 Roe, B. A., et al., *Biochem. biophys. Res. Commun.*, 66, 1097–1105 (1975).
37 Fink, L. M., Lanks, K. W., Goto, T., and Weinstein, I. B., *Biochemistry*, 10, 1873–1878 (1971).

to laser excitation of rhodopsin. Highly pure rhodopsin solutions in 0.067 M phosphate buffer (2% Ammonyx LO) were prepared as described previously[14–16].

Characteristic oscillograms, showing the main transient decay patterns after pulsed (530 nm, 40 ns) Nd laser excitation of rhodopsin solutions at 22 °C, are shown in Fig. 1. Oscillograms (*a–d* and *f*) were recorded using a pulse energy of ~20 mJ which is sufficiently high to reach saturation (absorbance changes independent of the pulse intensity), reflecting the attainment of a photostationary equilibrium between the primarily excited rhodopsin and photoproducts formed during the pulse (for example, PLR)[12]. Oscillogram *e* was recorded with a reduced pulse intensity, as indicated by the drop in the transient absorbance relative to, for example, *f*. Figure 1*a* and *b* shows that the initial changes at the end of the pulse are followed by decay and growing-in processes, exhibiting a half life of ~100 ns. These fast events are followed by much slower changes in the μs range (Fig. 1*c* and *d*). The corresponding difference spectra, recorded 50 ns, 700 ns and 1 ms from the origin of the pulse (Fig. 2), are proportional to the differences $D_T^\lambda - D_R^\lambda$, where D_R^λ is the drop in the absorbance of rhodopsin due to photobleaching and D_T^λ is the (time-dependent) absorbance of photoproducts present after the pulse. The dotted curves in the range where experimental points are not available were estimated as described below.

Primary processes in photochemistry of rhodopsin at room temperature

MOST of the information concerning the primary photochemical events in vision has been derived from low temperature work[1–5]. Recently, pulsed laser excitation methods with nanosecond[6–8] or picosecond[9] time resolution have been applied, in an attempt at clarifying the primary act at physiological temperatures. These studies showed that prelumirhodopsin PLR; or bathorhodopsin), which is the primary photoproduct at 77 K, is also formed at room temperature, with a rise time of less than 6×10^{-12} s (ref. 9) and a half life of ~10^{-7} s (refs 6–9). The problem remained, however, as to whether other primary transients, in addition to PLR, are formed at room temperature[8]. It is also important to establish if PLR is the unique precursor of the visual sequence at physiological temperatures, ruling out direct routes to subsequent intermediates—for example, lumirhodopsin (LR)—circumventing PLR. The present pulsed Nd laser study has attempted to clarify the above points, essentially testing the applicability of the low temperature photochemical picture (refs 10 and 11, and T.R., B. Honig, M.O. and T. Ebrey, unpublished) in physiological conditions.

The Nd laser photolysis system has been described previously[12]. Photolysis of the rhodopsin samples by the monitoring light was minimised by using short (10^{-2} s) exposures (obtained with a synchronised camera shutter) and appropriate combinations of Corning glass filters. Nanosecond measurements (with the monitoring source used in an intense pulsed mode) were not carried out in the 480–520-nm range, where no such filter combinations could be found, to reduce bleaching by the monitoring light (to less than ~5%) without lowering its intensity (*I*) below the minimum required[13]. The range between 520 and 550 nm was also avoided because of interference from light scattered by the laser. The above limitations on *I* impose serious restrictions on the signal-to-noise ratio (S/N) of the present nanosecond laser experiments (Fig. 1*a* and *b*, S/N = 5–10). A more intense monitoring pulse, however, significantly improving the accuracy of the measurements, pumps the system to an undefined mixture of rhodopsin and its bleaching products, preventing a quantitative analysis of the net effects due

Fig. 1 Characteristic oscillograms in the Nd laser photolysis of rhodopsin at room temperature. The initial absorbance of the solutions at 495 nm ranged from 0.5 to 1.0. Experiments *a*, *b*, *c* and *f* were carried out using a 20-mJ pulse, whereas *e* was carried out using a reduced intensity (oscillator only, without amplification). Load resistors used: 50Ω (*a*, *b* and *d*); 10 KΩ (*c*) and 100 KΩ (*e* and *f*). 100% (*c–f*) and 2% (*f*) denote the percentage of transmitted monitoring light. Upper traces in *c–f* were recorded in the absence of the monitoring beam. Vertical sensitivity 20 mV (*a*, *b*, *d* and *e*, central trace); 10 mV (*c*); 200 mV (*e*, bottom and upper traces); 100 mV (*f*). The initial light to dark deflections (not shown) in *a* and *b* were 180 mV and 170 mV, respectively. The break in the lower traces in *c* and *e* is due to firing of the laser pulse. In all other cases the laser was fired almost simultaneously with the oscilloscope trigger. In oscillogram *f* the lower trace denotes the transmittance of a bleached rhodopsin solution. The two central traces were recorded in an unbleached solution, without (upper) and with (lower) the laser being fired. From this information it is possible to evaluate the absorbance of the solution before pulsing and the change induced by the laser pulse between 1 and 7 ms. This enables evaluation of percentage rhodopsin bleached, as described in the text.
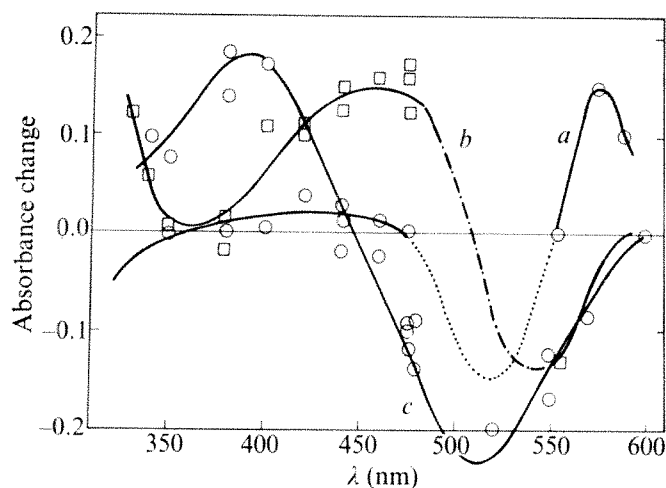
Fig. 2 Transient absorbance changes after pulsed Nd laser excitation of rhodopsin at room temperature. Data were recorded 50 ns (*a*), 700 ns (*b*) and 1 ms (*c*) after firing the laser. In *a* and *b* the spectral resolution is of the order of 5 nm (above 350 nm) or 10 nm (below 350 nm).

Oscillograms *c* and *d* in Fig. 1 and curve *c* in Fig. 2 show that after ~ 0.3 ms, bleaching in the range of the 500-nm rhodopsin band is associated with an increase in absorbance at about 380 nm, characteristic of the spectrum of meta-rhodopsin II (MII)[2,17,18]. Similarly, the difference spectra *a* and *b* may be associated with the generation of PLR (λ_{max} at 77 K = 543 nm) and LR (λ_{max} = 490 nm). (In the present experiments we have not attempted to identify MI, the precursor of MII, the absorption spectrum of which is similar to that of LR.) A closer examination of Fig. 2 indicates, however, that the observed difference spectra cannot be rationalised only in terms of PLR, LR and MII. In all three cases the absorbance change in the 450–550-nm range is higher than that expected from the corresponding low temperature differences in extinction coefficients[1,17,18]: $\varepsilon_{M_{II}}^{\lambda} - \varepsilon_L^{\lambda}$, $\varepsilon_R^{\lambda} - \varepsilon_R^{\lambda}$ and $\varepsilon_{PLR}^{\lambda} - \varepsilon_R^{\lambda}$, indicating that apart from PLR an additional product (X) is present at the end of the pulse which is stable at least up to the stage of MII. A new band in the neighbourhood of 450 nm was also reported by Bensasson *et al.*[8] along with a peak at about 330 nm, not observed in the present work.

Information relevant to the identification of X was obtained by measuring its contribution, reflected by the absorbance drop at 480 nm (at the stage of MII) relative to that at 550 nm, as a function of the laser pulse intensity. Excitation below saturation (for example, Fig. 1*e*), when the amount of light absorption by PLR relative to rhodopsin is considerably reduced, leads to an increase in the above ratio ($D_{480}/D_{550} = 2.2 \pm 0.4$), compared with that ($0.9 \pm 0.3$) obtained using saturating excitation energies (for example, Fig. 2*c*). It seems therefore that X is a secondary product, formed during the pulse because of light absorption by PLR. Thus, in view of the photoequilibrium

$$\text{rhodopsin} \underset{h\nu}{\overset{h\nu}{\rightleftarrows}} \text{PLR} \underset{h\nu}{\overset{h\nu}{\rightleftarrows}} \text{Isorhodopsin (IR)}$$
$$(\lambda_{max} \text{ at } 77 \text{ K} \qquad (\lambda_{max} \text{ at } 77 \text{ K} \qquad (\lambda_{max} \text{ at } 77 \text{ K}$$
$$= 498 \text{ nm}) \qquad\qquad = 543 \text{ nm}) \qquad\qquad = 491 \text{ nm})$$
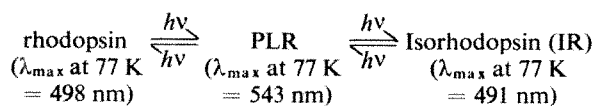
which is known to be established at 77 K (refs 3 and 19), it is strongly suggested that X should be identified with the 9-*cis* isomer, IR (λ_{max} at 25 °C = 485 nm)[11].

This suggestion may be more quantitatively tested by considering the change in absorbance after 1 ms at 550 nm (where $\varepsilon_{MII} \sim 0$)

$$\Delta D_{550} = C_{IR}\varepsilon_{IR}^{550} - C_{MII}\varepsilon_R^{550} - C_{IR}\varepsilon_R^{550}$$

Taking[11] $\varepsilon_{IR}^{550}/\varepsilon_R^{550} \sim 0.5$ and neglecting the contributions of

IR and rhodopsin at 380 mm (that is, approximating: $\Delta D_{380} \sim \varepsilon_{MII}^{380} C_{MII}$), we have

$$\Delta D_{550} = -\varepsilon_R^{550}(C_{IR}/2 + \Delta D_{380}/\varepsilon_{MII}^{380})$$

From the data of Fig. 2 (with $\varepsilon_{MII}^{380} = 4.0 \times 10^4$ M^{-1}cm^{-1} and $\varepsilon_R^{550} = 1.8 \times 10^4$ M^{-1}cm^{-1}) we obtain $C_{MII} = 4.5 \times 10^{-6}$ M and $C_{IR} = 6.6 \times 10^{-6}$ M. The concentration of bleached rhodopsin is $C_{IR} + C_{MII}$ and its contribution to the difference spectra in Fig. 2, $D_R^{\lambda} = \varepsilon_R^{\lambda}(C_{IR} + C_{MII})$ is shown in Fig. 3*a*. Addition of D_R^{λ} to each of the difference spectra in Fig. 2 yields curves *d*, *e* and *f* in Fig. 3, representing correspondingly D_T^{λ} (50 ns), D_T^{λ} (700 ns) and D_T^{λ} (1 ms). It can be seen that curve *f* can be represented by the superimposition of the known spectra[11,17,18] of MII (Fig. 3*c*) and IR (Fig. 3*b*), in keeping with the identification of X as IR. A similar procedure shows that curves *e* and *d* are obtainable by superimposing the contribution of IR and correspondingly (allowing for a 10–20-nm temperature shift and a $\sim 10\%$ change in extinction coefficients) those of LR and PLR in amounts equal to that of MII.
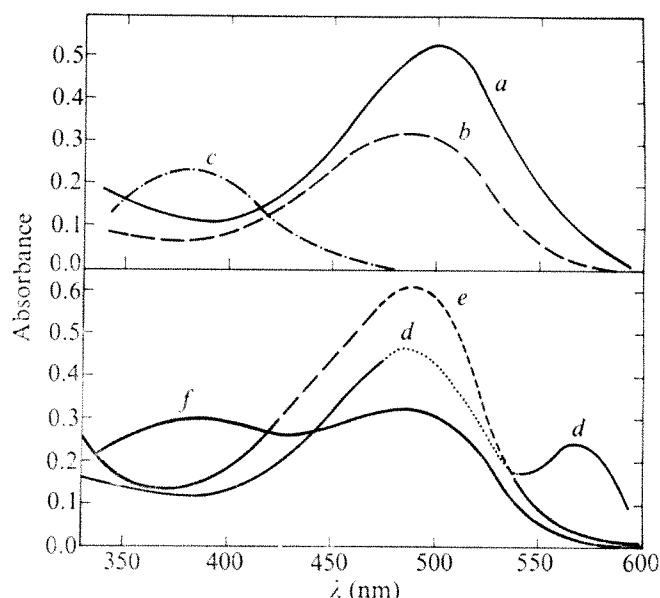
Fig. 3 *a*, Contribution of bleached rhopodsin (D_R^{λ}) to the difference spectra of Fig. 2; *b* and *c*, spectra of MII and of transient X (IR), respectively. *d–f*, Spectra of transients present after 50 ns (*d*), 700 ns (*e*), and 1 ms (*f*) evaluated from the corresponding traces in Fig. 2, as described in the text.

The above observations are consistent with PLR being the unique primary product in the photochemistry of rhodopsin at room temperature. It is also evident that it constitutes the precursor of the subsequent intermediates in the cycle (LR, MI, MII, and so on). Thus, the classical sequence of primary events as derived for rhodopsin on the basis of low temperature work[1-5] seems to be applicable at physiological temperatures, in agreement with a conformational (for example, *cis-trans*) change as the only primary step of the visual cycle (refs 10 and 11, and T.R. *et al.*, unpublished).

We thank Dr B. Honig, Dr A. Kropf and Dr R. Bensasson for discussions. T.R. is grateful for Professor T. Ebrey's invitation to join his laboratory in the State University of Illinois.

CH. R. GOLDSCHMIDT
M. OTTOLENGHI
T. ROSENFELD

Department of Physical Chemistry,
Hebrew University,
Jerusalem, Israel

1 Wald, G., *Science*, **162**, 230 (1968).
2 Abrahamson, E. W., and Wiesenfeld, J. R., in *Photochemistry of Vision* (edit. by Dartnall, H. J. A.), 69–121 (Springer, Berlin, Heidelberg and New York, 1972).
3 Yoshizawa, T., *ibid.*, 146–179.
4 Grellman, K. H., Livingston, R., and Pratt, D., *Nature*, **193**, 1258 (1962).
5 Abrahamson, E. W., in *Biochemistry and Physiology of Visual Pigments*, (edit. by Langer, H.,), 47–56 (Springer, Berlin, Heidelberg and New York, 1973).
6 Cone, R. A., *Nature new Biol.*, **236**, 39 (1972).
7 Rosenfeld, T., Alchalal, A., and Ottolenghi, M., *Nature*, **240**, 482 (1972).
8 Bensasson, R., Land, E. J., and Truscott, T. J., *Nature*, **258**, 768 (1975).
9 Busch, G. E., Applebury, M. L., Lamola, A. A., and Rentzepis, P. M., *Proc. natn. Acad. Sci. U.S.A.*, **69**, 2802 (1972).
10 Kropf, A., and Hubbard, R., *Ann. N.Y. Acad. Sci.*, **74**, 266 (1958).
11 Yoshizawa, T., and Wald, G., *Nature*, **197**, 1279 (1963).
12 Goldschmidt, C. R., Ottolenghi, M., and Korenstein, R., *Biophys. J.* (in the press).
13 Goldschmidt, C. R., in *Lasers in Physical Chemistry and Biophysics* (edit. by Joussot-Dubien, J.), 499 (Elsevier, Amsterdam, 1975).
14 Papermaster, D. S., and Dreyer, W. J., *Biochemistry* **13**, 2438 (1974).
15 Shichi, H., *Expl Eye Res.*, **17**, 533 (1973).
16 Ebrey, T., *Vision Res.*, **11**, 1007 (1971).
17 Matthews, R., Hubbard, R., Brown, P. K., and Wald, G., *J. gen. Physiol.*, **47**, 215 (1963).
18 Yoshizawa, T., and Horiuchi, S., *Biochemistry and Physiology of Visual Pigments* (edit. by Langer, H.), 69–81 (Springer, Berlin. Heidelberg and New York, 1973)
19 Yoshizawa, T., and Wald, G., *Nature*, **197**, 1279 (1963).

Structure of thymidylyl-3',5'-deoxyadenosine

THE triethylammonium salt of the dinucleoside phosphate, thymidylyl-3',5'-deoxyadenosine (TpdA) forms helical structures when aqueous solutions are allowed to evaporate[1]. There are some similarities between the X-ray diffraction patterns from TpdA and those from poly(dA-dT)–poly(dA-dT), and we originally thought that the structures might be similar[1]. But when the structure for poly(dA-dT)–poly(dA-dT) was proposed[2] it was apparent that the TpdA structure was not of the same kind. This, together with the fact that we had obtained more detailed X-ray diffraction patterns from better oriented TpdA specimens, led us to consider the different molecular model for TpdA which we describe here.

Figure 1a shows the diffraction pattern from a specimen of TpdA at ~ 60% relative humidity. The layer-line spacing corresponds to a periodicity of 26.4 Å and the equatorial Bragg reflections fit a hexagonal lattice with $a = 21.3$ Å. Discrete reflections are confined to the inner region of the equator—the remainder of the pattern shows continuous intensity along the layer lines. The strongest intensity is in the meridional region of the 7th layer line. The intensity on the 8th layer line is also very strong as well as the first maximum on the 1st layer line. The near meridional diffraction on layer lines 5–8 is similar to the diffraction in the same region in patterns from DNA, double-helical RNA and synthetic polynucleotides. It is mainly due to the bases and gives information about their tilt relative to the helix axis.

In the early photographs it was difficult to decide on which of the layer lines 5–8 the intensity was truly meridional and thus to decide the number of residues per turn of the helix. Better oriented specimens, tilted at 12° from perpendicular to the X-ray beam show that the intensity on the 5th and 8th layer lines is not meridional (Fig. 1b). But it is not possible to decide whether the truly meridional intensity occurs on the 6th or the 7th layer lines. Tilting a specimen at 24° from perpendicular to the X-ray beam to examine the meridional region of the 12th and 14th layer lines, it is again difficult to decide which of these two has truly meridional intensity (Fig. 1c). In fact the intensity on the 13th layer line seems to be more meridional, although this appearance could be a result of the intensity distribution on these layer lines, which is very dependent on the tilt of the bases.

Meridional intensity on the 7th and 14th layer lines would imply seven residues per turn of the helix, with a rise per residue of 3.77 Å and a rotation per residue of 51.4°. Meridional intensity on the 6th and 12th layer lines would imply six residues per turn, with corresponding translational and rotational values of 4.4 Å and 60°.

Our original idea that the TpdA structure might be similar to that of poly(dA-dT)–poly(dA-dT) would mean that the repeating unit of the structure was a pair of TpdA molecules related to each other by a twofold rotation axis perpendicular to the helix axis. The large tilt of the bases required to give the appropriate rise per residue in such a model is not however, consistent with the near-meridional intensity distribution in the diffraction pattern. We therefore considered alternative models, in which the repeating unit is one TpdA molecule and with a smaller tilt of the bases, as suggested by the near-meridional diffraction intensity. We assumed that within the dinucleoside phosphate, the two nucleosides are related by a right-handed screw, as in all other helical polynucleotide structures. We find that it is possible to construct a stereochemically acceptable model with seven TpdA molecules per turn of the helix, and which also gives a calculated Fourier transform which is in reasonably good agreement with the observed diffraction pattern (Fig. 2). A similar sixfold model does not give good agreement with observation.

The thymine of one TpdA molecule is hydrogen bonded to, and coplanar with, the adenine base of the next. This is only possible if one TpdA molecule is related to the next by a left-handed screw. The model is thus a left-handed helix, although within the dinucleoside phosphate unit, the nucleosides are related as in a right-handed helix. To use protein structure terminology, the secondary structure is right handed but the quaternary structure is left handed. Adenine and thymine are
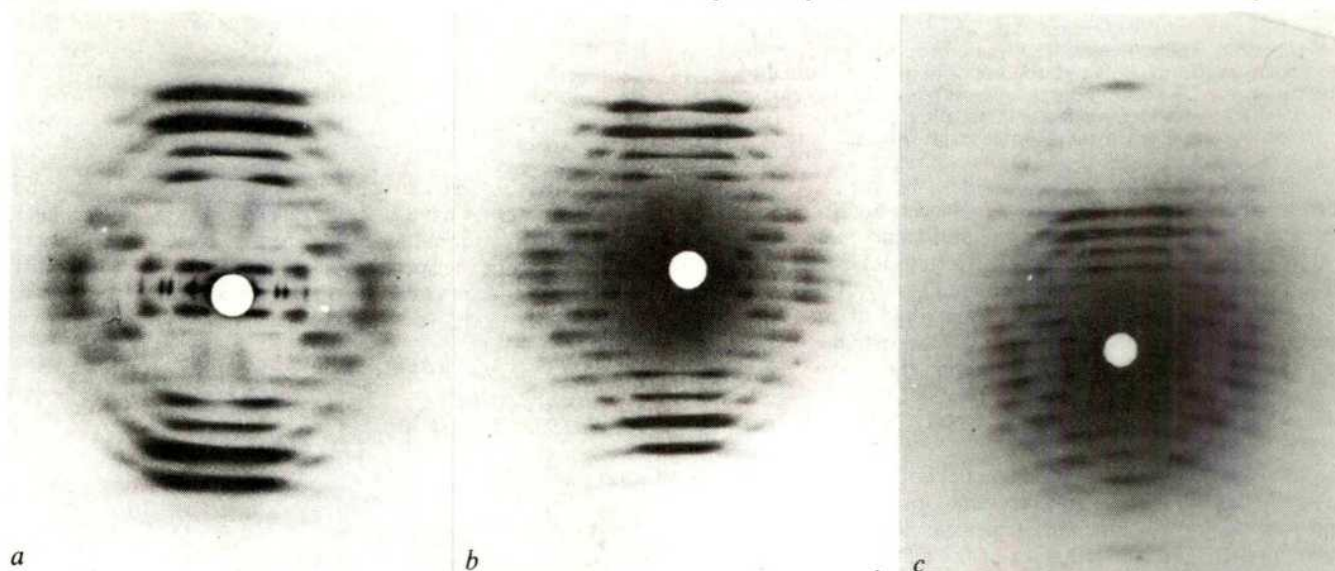
Fig. 1 X-ray diffraction patterns from TpdA specimens at ~ 60% relative humidity. Diffraction photographs were taken on a North American Phillips microcamera, using a Hilger and Watts semimicrofocus X-ray generator. *a*, Specimen approximately perpendicular to X-ray beam; *b*, specimen tilted through 12°; *c*, specimen tilted through 24°.
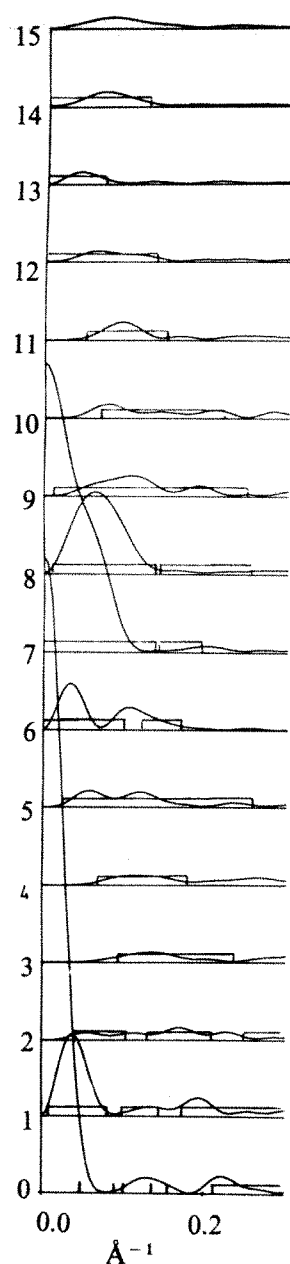
Fig. 2 The cylindrically averaged molecular Fourier transform of the sevenfold TpdA model. The smooth curves indicate the calculated transform. Bragg reflection positions are indicated by vertical bars and the continuous intensity is represented by blocks. Modified atomic scattering factors[8], to allow for water in the structure, were used in the Fourier transform calculations.

Fig. 3 Projections of one TpdA molecule together with the adenine base of the next molecule above and the thymine base of the next molecule below. *a*, View perpendicular to the helix axis; *b*, view along the helix axis.

observed in some nucleotides[5-7]. A *gauche–gauche* conformation about the C(4′)–C(5′) bond consistently gave rise to an unacceptably short contact. Two views of the model are shown in Fig. 3, and a sketch of the model is shown in Fig. 4. There is a large degree of overlap between adenine bases but very little between

Fig. 4 Schematic representation of the model.

linked to each other by Hoogsteen[3] hydrogen bonds, and both adenine and thymine bases are in the anti conformation about the glycosyl bonds. The interaction between bases is thus similar to that between chains 2 and 3 in poly(dT)–poly(dA)–poly(dT) (ref. 4). The adenine–thymine bases are tilted about 9° from perpendicular to the helix axis so that the distance between adjacent adenine–thymine bases is less than the rise per residue. The glycosyl torsion angle O(1′)–C(1′)–N(9)–C(4) is 176° for the adenine base and the angle O(1′)–C(1′)–N(1)–C(2) is 220° for the thymine.

Both furanose rings have C3′ endo puckers, so the conformation is similar to that in A-DNA and double-helical RNA. Sugar rings with other kinds of pucker did not give such satisfactory models. The conformation of the C5′–O5′ bond about the C4′–C5′ bond in the link between the nucleosides is in the *gauche–trans* region. Although this is not the preferred conformation for this bond, similar conformations have been

the thymine bases. The structure is polar and the fact that the diffraction patterns have discrete reflections only in the inner region of the equator indicates that the 'up' and 'down' arrangement of helices in the specimens is random.

We thank Mr Ron Damaschka for computational assistance.

H. R. WILSON
JENAN AL-MUKHTAR

Carnegie Laboratory of Physics,
University of Dundee,
Dundee DD1 4HN, UK

1 Wilson, H. R., *Nature*, **251**, 735–736 (1974).
2 Arnott, S., Chandrasekaran, R., Hukins, D. W. L., Smith, P. J. C., and Watts, L., *J. molec. Biol.*, **88**, 523–533 (1974).
3 Hoogsteen, K., *Acta Cryst.*, **16**, 907–916 (1963).
4 Arnott, S., and Selsing, E., *J. molec. Biol.*, **88**, 509–521 (1974).
5 Saenger, W., and Suck, D., *Nature*, **242**, 610–614 (1973).
6 Young, D. W., Tollin, P., and Wilson, H. R., *Nature*, **248**, 513–514 (1974).
7 Viswamitra, M. A., and Seshadri, T. P., *Nature*, **252**, 176–177 (1974).
8 Langridge, R., *et al.*, *J. molec. Biol.*, **2**, 38–64 (1960).

matters arising

Energy expenditure in children

IT was with interest that we read the paper by Griffiths and Payne[1], in which they presented interesting data and then discussed the possible relevance of environmental and genetic factors on the level and nature of the energy balance obtained with the two groups of children.

It is interesting to note that on expressing the mean expenditure and intake data for both groups as a ratio × 100, precisely the same figure ensues

$$\frac{\text{Energy expenditure}}{\text{Energy intake}} \times 100$$

Non-obese	Obese
$\frac{1,508}{1,433} \times 100 =$	$\frac{1,174}{1,115} \times 100 =$
105.23	105.29

Is it not possible therefore, that few overall metabolic differences exist between the 5-yr-old children of these two groups of parents, which by implication, suggests the predominance of environmental factors on these results?

It is surprising that the O group children were found to have lower daily energy intakes than their controls (N group). Therefore it would seem likely that parental pride during the nutritional survey may have had a marked influence.

We are also a little concerned that no comment was made about the data which showed that both groups of children were in negative energy balance, which for 5 yr olds would seem to be unlikely. From our experience SAMI measurements can be misleadingly high unless very precise and immobile electrical contacts are made, and thus background noise is eliminated.

B. G. MILLER
W. R. OTTO

*Department of Physiology
and Biochemistry,
University of Southampton,
Southampton SO9 3TU, UK*

1 Griffiths, M., and Payne, P. R., *Nature*, 260, 698–700 (1976).

GRIFFITHS AND PAYNE REPLY—In our view the most important findings in our study were the differences in total daily energy metabolism, and of resting metabolic rate between the two groups of children.

We agree that the measurements of daily expenditure by the SAMI technique are inherently less reliable than are those of food energy intake, and have offered them as supporting evidence. That is to say, unlike Miller and Otto[1], we regard the equal ratios of expenditure to intake as contributing positively to the evidence for different amounts of energy metabolised daily by the two groups.

Because the differences between the estimations of expenditure and intake (energy balance) are small in relation to the errors inherent in the two measurements, the energy balances do not statistically differ significantly from zero. We are not therefore inclined to attach any biological significance to the fact that they happen, on average, to be negative in both groups.

It is of course possible that one aspect of the behavioural differences between obese and normal people is that the former tend to 'cheat' when attempts are made to measure habitual intake, either their own or their children's. We can only repeat that: the O group children were not themselves measurably overweight so there was no obvious reason for the parents to take pride in suggesting that they eat sparingly, the degree of cooperation and the level of communication with the subjects about the objectives of the study were of necessity, very high, the admittedly less reliable estimates of expenditure closely parallel those of intakes, and show the same magnitude of difference between the groups.

*London School of Hygiene
and Tropical Medicine*

1 Miller, B. G., and Otto, W. R., *Nature*, 263, 173 (1976).

Benzodiazepines and GABA

BASED on a few experiments on Deiters' neurones and cerebellar Purkinje cells, Steiner and Felix[1] concluded that benzodiazepines antagonise the inhibitory effect of GABA. To understand the mode of action of centrally acting agents, the evidence available from various pharmacological experiments must be considered and the manifold pitfalls inherent in the single-cell recording and microelectrophoretic techniques taken into account. The conclusion of Steiner and Felix[1] is in contrast to the well documented facts that benzodiazepines: (1) are the most potent antagonists of convulsions induced by picrotoxin and bicuculline[2-4] (generally accepted to block GABA receptors); (2) enhance presynaptic inhibition in the spinal cord and in the cuneate nucleus[2,5,6]; (3) enhance postsynaptic inhibition in the cuneate nucleus[6] and in the substantia nigra[2,7] and (4) are mutual antagonists of GABA receptor blockers at inhibitory synaptic processes believed to be mediated by GABA[5-7]. We think that the conclusion of Steiner and Felix[1] is based on the erroneous interpretation of experimental conditions that were inadequate for the specific problem.

In their experiments on vestibular neurones, they made the implicit assumption that the pulses applied to the cerebellar white substance (in their Fig. 1 either the time calibration is wrong or the rate of stimulation is not that indicated in the legend) released the same amount of GABA within the Deiters' nucleus before and after diazepam. This crucial premise is probably not correct for the following reasons. The stimulating electrode excites Purkinje cell axons directly and, in addition, evokes an unknown number of synaptically evoked discharges of Purkinje cells through the inevitable stimulation of cerebellar afferents. The number of the latter discharges depends on the balance between mono- and disynaptic excitation and the inhibitory influence on Purkinje cells. If diazepam, as we assume, enhances GABA-mediated synaptic inhibition also in the cerebellar cortex, it should reduce the number of Purkinje cell discharges evoked synaptically by the stimulus train, and thus, produce an apparent reduction of the inhibitory effect on Deiters' neurones, which was interpreted by Steiner and Felix[1] as an antagonism of the effect of GABA released within the vestibular nucleus. The apparent reduction by diazepam of the inhibitory effect of iontophoretically applied GABA may be explained in the same way. In the pre-drug period, the depression of Deiters' neurones is achieved by the fixed amount of

exogenous GABA which adds to the endogenous GABA that is released by the continuous activity of Purkinje cells. After diazepam, the spontaneous firing rate of Purkinje cells is markedly reduced[2,8]. This must result in disinhibition of Deiters' neurones and in the reduction of the total amount of GABA (exogenous plus endogenous) acting on them. The validity of our reasoning could easily be checked experimentally.

Any interpretation of the results of Steiner and Felix[1] on cerebellar Purkinje cells must remain highly speculative when taking into consideration that four of the five types of cerebellar cortical neurones are believed to be GABA-ergic and that the activity of the Purkinje cell is itself inhibited by GABA-ergic interneurones. In the experiments of Steiner and Felix[1] the benzodiazepines did not seem to depress the firing rate of Purkinje cell, whereas in our own experiments[2,8] several representatives of this class of compounds regularly and markedly reduced the spontaneous discharges and strongly antagonised the opposite effect of bicuculline. We should also like to emphasise that the experiments of Steiner and Felix[1] were performed on anaesthetised animals and a multitude of interactions between anaesthetics and benzodiazepines have been described.

In conclusion, it seems that the findings of Steiner and Felix[1] do not fail to support a facilitating effect of benzodiazepines on GABA-mediated transmission[2] but rather that they are easily explained by such a mode of action.

W. Haefely
L. Pieri
P. Polc
R. Schaffner

Pharmaceutical Research Division,
F. Hoffmann–LaRoche & Co. Ltd,
Basle, Switzerland

[1] Steiner, F. A., and Felix, D., *Nature*, **260**, 346–34 (1976).
[2] Haefely, W., *et al.*, in *Mechanisms of Action of Benzodiazepines* (edit. by Costa, E., and Greengard, P.), 45–61 (Raven, New York, 1975).
[3] Costa, E., *et al.*, *Life Sci.*, **17**, 167–186 (1975).
[4] Curtis, D. R., Game, C. J. A., and Lodge, D., *Br. J. Pharmac.*, **56**, 307–311 (1976).
[5] Polc, P. Möhler, H., and Haefely, W., *Archs Pharmac.*, **284**, 319–337 (1974).
[6] Polc, P., and Haefely, W., *Archs Pharmac.* (in the press).
[7] Keller, H. H., Schaffner, R., and Haefely, W., *Archs Pharmac.* (in the press).
[8] Pieri, L., *Experientia* (in the press).

STEINER AND FELIX REPLY—The hypothesis advanced by Haefely *et al.*[1] that benzodiazepines facilitate GABA-responsive neurotransmission is based on the premise that GABA is the proven mediator of inhibition at presynaptic sites (this is still controversial, see refs 2 and 3) as well as on the demonstration

---

that the action of the presumptive GABA antagonists bicuculline and picrotoxin is reduced by this class of psychotropic drugs. Haefely *et al.*[1] are justified in pointing out that the mode of action of centrally acting drugs is bound to be complex and that more than one level of interaction ought to be considered. They fail to mention, however, that so far, there is no direct experimental support for their contention at any level of organisation tested: Neither their own work, nor that of others cited in support of their view, has directly tested GABA-mediated transmission, and thus all their experimental evidence is circumstantial. It was precisely for this reason that we chose to test their hypothesis at the level of identified Purkinje cells and Deiters' neurones[4]. At both these sites, there is incontrovertible evidence for GABA-mediated inhibition, the extent of which can readily be followed by standard electrophysiological techniques. In this anatomically and functionally well defined system, our results indicated that benzodiazepines inhibit GABA-responsive transmission[4]. We remain convinced that our technique is adequate to study the problem, and welcome this opportunity to answer questions raised by Haefely *et al.*[1] about our approach.

For the orthodromic activation of Deiters' neurones by stimulation of Purkinje cell axons, we deliberately chose supramaximal parameters (trains of three pulses of 250 μs duration at 500 -μs intervals and an intensity of 0.8 mA). Although we do not refute the view expressed[1] that this type of stimulations is likely to have repercussions on the cerebellar cortex as well as on nerve endings at the level of the Deiters' nucleus, we believe that, as in other biological systems, orthodromically stimulated transmitter release is a direct function of the stimulus applied (and therefore, in our case, a constant); we cannot accept their view that the amount of GABA

---

released in our experimental conditions is modified, in any important respect, by way of the soma of the Purkinje cells itself. With respect to the methodology used, we concede that in any study using microelectrophoresis, the quantity of exogenous neurotransmitter administered, albeit small in absolute terms, is likely to be in the pharmacological range. Changes in endogenous GABA release would therefore seem negligible compared with the amount of exogenous GABA administered. Haefely *et al.*[1] argue that a drop in endogenous GABA release by Purkinje cells (which, in keeping with their hypothesis, they ascribe to a facilitation of GABA-mediated inhibition by benzodiazepines at this site) could "easily" be distinguished from the fixed amount of exogenous GABA applied; in our view, this would require the direct monitoring of GABA release at Purkinje cell terminals on Deiters' neurones and could not reliably be deduced from the rate of cell firing alone (to our knowledge, such microtechniques are presently beyond our reach).

If Haefely *et al.* were correct in their assumption that benzodiazepines facilitate GABA-mediated inhibition at the Purkinje cell (and so reduce their GABA release), they would also have to concede a similar synergism at the GABA-responsive synapses impinging on the Deiters' neurones themselves; if their argument is taken to its logical conclusion, one would have to postulate that reduced GABA release at this site is counteracted, and possibly cancelled out, by an increase in its biological effect. This assumption, however is not borne out by our findings. Likewise, we consider the absence of benzodiazepine effects on the spontaneous rate of discharge of Purkinje cells as further evidence that our findings are not secondary to changes in the extent of excitation of the cerebellar cortex or the use of anaesthetised animals but are a reflection of drug interaction occurring at the site of study itself.

Our results, as well as those of Gähwiler[5], obtained in simple well defined systems, indicate a functional antagonism between GABA and benzodiazepines. They do not confirm the notion of a synergism inferred from indirect observations on more complex structures.

Institute for Brain Research,
University of Zurich,
8029 Zurich, Switzerland

[1] Haefely, W., *et al.*, *Nature*, **263**, 173–174 (1976).
[2] Ryall, R. W., *Amino Acid Neurotransmitters*, **4** (edit. by Iversen, L. L., Iversen, S. D., and Snyder, S. H.), 101–109 (Plenum, New York, 1975).
[3] Schmidt, R. F., *Ergebn. Physiol.* **63**, 20–101 (1971).
[4] Steiner, F. A., and Felix, D., *Nature*, **260**, 346–347 (1976).
[5] Gähwiler, B. H., *Brain Res.*, **107**, 176–179 (1976).

reviews

Marking time

Colin Renfrew

Radiocarbon: Calibration and Prehistory. Edited by Trevor Watkins. Pp. vi+147. (Edinburgh University: Edinburgh, January 1976.) £3.50.

THE tree-ring calibration of radiocarbon dating has had a number of striking archaeological consequences, bringing the historical and radiocarbon calendars for the ancient civilisations of the Old World into closer agreement and pushing the date of some cultures and monuments back much earlier than had been realised.

The intention of the symposium from which this volume springs was to "allow discussion of the subject of correction before a lay audience in terms which they could understand". Most of the contributions succeed well in this modest task. Richard Burleigh, for instance, advocates preliminary calibration for many purposes, while pointing out that for other purposes uncalibrated dates can often legitimately be compared. This view, however, is somewhat over-simplified by the editor: "it is too early to use the curves at present being bandied around as calibration curves for radiocarbon dates". Unfortunately the volume went to press before the publication by R. M. Clark of the most statistically sound and reliable of these (*Antiquity*, 49, 252–272, 1975).

The volume as a whole betrays a similar lack of balance. Alongside useful reviews of European dendrochronology by Fletcher and of the problem of irregularities ('kinks') by Ottaway and Ottaway, is a substantial disquisition by McKerrell (53 pages long, more than a third of the total) arguing the very special case that the existing calibration curves are reasonably sound except for the period between 2,500 BC and 1,500 BC (in calendar years).

This view is based on the assessment of radiocarbon and historical dates (the later derived from the Egyptian calendar) for Egypt, the Near East and the Aegean. Unfortunately the archaeology for the Agean is distinctly shaky (p61) and the treatment of the central European material (Veterov, Madarovce, and so on; p68) might have been more convincing at the hand of a professional archaeologist.

It is surprising that a scientific worker should propose, in the proceedings of a conference designed for laymen, a theory with major geophysical implications which might better first be offered to the critical scrutiny of his physicist colleagues. Moreover, when the theory that calibrated dates in general are sound (with the specific exception of a single millennium), has global implications, would it not be wise to consider the broader consequences outside of Europe?

What begins as a useful introductory handbook rapidly becomes a specialised and highly controversial critique of calibration, an argument *à thèse* rather than a clarification, on which the comments of competent radiocarbon specialists might have been invited. The editor could usefully have asked himself whether he wanted his book to be a searching professional critique of calibration, or a lucid, non-technical exposition. There is still a need for both. ☐

Colin Renfrew is Professor of Archaeology at the University of Southampton, UK.

Textbook of ichnology

The Study of Trace Fossils: A Synthesis of Principles, Problems and Procedures in Ichnology. Edited by Robert W. Frey. Pp. xiv+562. (Springer: Berlin and New York, 1975.) DM 141.70; $58.10.

MANY sedimentary rocks contain fossilised branching structures that a hundred years ago were mistakenly identified as seaweeds and called 'fucoids'. In the 1880's these were shown to be the tracks of animals, and in most parts of the world their study lapsed. But in Germany this was not at all the case. In 1925 Rudolf Richter founded a new school of *Akutuopaläontologie* in which the traces of animals played a vital part in the determination of environment. These early studies led to the establishment of a new and exciting discipline in which palaeontologists were able to postulate behavioural parameters. In this way 'palaeoethology' came to join 'palaeoecology' as a branch of 'geobiology'. In these studies the Germans remained far in advance of the rest of the world for over thirty years; and, indeed, it can still be said that they lead the field, with Professor Adolf Seilacher of Tübingen (who contributes a foreword to this book) as the acknowledged dean.

But at least nowadays the supremacy of our German colleagues in this field has been challenged. The discipline of palaeontology has much changed from the time when it was dominated by morphology and taxonomy. Its boundaries have extended, and in 1970 Crimes and Harper edited a remarkably successful symposium on trace fossils which reported the proceedings of an international symposium held in Liverpool. This was followed in 1971 by a field-guide to trace fossils produced by Louisiana State University. It has now been realised all over the world that ichnology plays a quite indispensable role in field studies aimed at interpreting past environments.

The present volume is the only textbook on the subject available in English. It captures well the excitement of a subject which to most of the world is still seen as new. The editor has kept his contributors well disciplined, so that in spite of the fact that it combines the work of no fewer than 28 authors, it retains an admirable cohesiveness as well as being comprehensive. For example, it includes chapters on both plant traces and vertebrate trails (both by William Sarjeant), and concludes with a fascinating section describing two new developments in ichnology— one on experimental methods (By Christopher Elders) and the other describing George Farrow's brilliant techniques for the study of the traces of living animals in present day sediments.

This book will certainly be regarded as an indispensable teaching aid in all future courses dealing with ichnology or palaeoethology, and the editor, Robert Frey, is to be congratulated on the clarity of exposition he has been able to command as well as the logic he

has imparted to the work as a whole. It is all the more to be regretted that the publishers, in spite of their reputation and their experience, have allowed the most appalling errors to appear in the printing of the illustrations. Five of the tables in the book have had all data in them omitted in printing. In a tardy attempt at correcting this error, the publishers have issued with the later copies offered for sale an eight-page 'errata sheet', but with incredible incompetence these have been printed in a way that makes it impossible to substitute them for the faulty pages, as each has been backed on to another, with p472 backing, for example, p516. It can only be hoped that a new printing can put such a stupid blunder right.

P. C. Sylvester-Bradley

P. C. Sylvester-Bradley is F. W. Bennett Professor of Geology in the University of Leicester, UK.

Reaction kinetics

The Organic Chemistry of Electrolyte Solutions. (Interscience Monographs on Organic Chemistry.) By John E. Gordon. Pp. xxi+554. (Wiley-Interscience: New York and London, December 1975.) £15.35.

The Electrolyte Chemistry of Organic Solutions? No, that will not do either; the title is strange and ill-chosen.

The book's aim, according to the author, is to provide the organic chemist with everything he needs to know about electrolyte solutions, to treat all important solution chemistry measurements in one of its three chapters, while avoiding treating in detail aspects accessible through books such as Robinson and Stokes, and Harned and Owen; and therefore to graft on to the traditional physicochemical framework more 'chemical' information lacking in the great books from non-traditional sources such as vibrational and nuclear magnetic resonance spectroscopy. In less pretentious phraseology, to up-date the standard texts by including the important advances over the past 20 years made through the application of spectroscopic techniques. This is not the first attempt to do so in spite of the fly-leaf's claim, nor does the book seem to have been produced any quicker than a multiauthored volume, as the Series Editor claims, for the literature coverage is complete only to mid-1973 and cuts off in August 1974; and the book was not published until the end of 1975.

The three chapters referred to above are called Salt Effects, Ion Solvation and Ion Association, which are not mutually exclusive categories, as

the author realises. Each chapter concludes with a lengthy section discussing applications to kinetic studies, thus revealing the true interest of the author and his intentions. Has he succeeded? To a very large measure, he has. The book is no paste and scissors concoction, but scholarly and nearly always sound in approach and explanation if not in balance and selection of material. For, in discussion of theoretical conductance equations, only the earlier work of Fuoss is discussed, neglecting his later contributions and those of Prue, Pitts and Justice amongst others. Similarly the discussion of *p*H, under the heading of 'Single ion activity coefficients', returns unnecessarily to the 1920s treatment; and the widely adopted important NBS method is described as "based on yet another convention". Order of introduction of new concepts could have been better conceived. Terms which will surely puzzle the uninitiated when they first appear often receive more detailed explanation many pages later. The book is well illustrated with figures and diagrams drawn from the literature but not always too well associated with textual explanation. Nevertheless, physical organic chemists and others should find much of value in this book as a source book to the extensive literature on electrolyte solutions. It should help to continue to advance the interpretation of reaction kinetics in organic solvents, a subject which has progressed a long way since ion-pairs and solvated ions were first invoked to explain kinetic data when mechanistic explanation was otherwise at a low ebb.   **A. K. Covington**

A. K. Covington is Reader in Physical Chemistry, University of Newcastle upon Tyne, UK.

Cell split

Cell Division in Higher Plants. (Experimental Botany: An International Series of Monographs, Vol. 7.) Edited by M. M. Yeoman. Pp. x+542. (Academic: London and New York, April 1976.) £16.50; $41.

KNOWLEDGE of the cell division process and its control in plants has developed rapidly in recent years from observations made in a diversity of research fields. Dr Yeoman's concept behind this volume was clearly that this scattered material had to be brought together and presented in a developing sequence so that this recent progress could be fully appreciated and the as yet outstanding problems critically identified. The Botany Department at the University of Edinburgh has been one of the main centres of current research on cell division in plants, and

members of this department have obviously played a critical role, under Dr Yeoman's astute editorship, in the realisation of this volume. Professor R. Brown, in an introductory chapter, identifies the problems posed by the organised cell divisions in plant meristems and develops, in a thought-provoking way, models which will clearly guide future research. A group of chapters then deal with the process of cell division both in somatic and reproductive cells. A. F. Dyer contributes two chapters on mitosis which deal in a most lucid way with mitotic cycles and the origin and consequences of mitotic modifications and errors. A parallel chapter by M. D. Bennett (Plant Breeding Institute, Cambridge) deals with meiosis. M. M. Yeoman and P. A. Aitchison summarise our present very limited knowledge of the molecular events of the cell cycle, and Rachel M. Leech (York) our similarly very imperfect understanding of plastid replication.

The second major section of the book considers cell division in relation to the generation of form. The very intensive studies on the root apex are surveyed by F. A. L. Clowes (Oxford). This is followed by chapters on the shoot apex (R. F. Lyndon), leaves (J. E. Dale) and cambium (I. D. J. Phillips, Exeter). W. A. Jensen (Berkeley, California) contributes a disappointing short chapter on cell division in embryo development. This section of the book is completed by an interesting survey of unorganised cell division and morphogenesis in tissue and cell cultures by A. W. Davidson, P. A. Aitchison and M. M. Yeoman.

A final section headed Summary and Perspectives and written by Dr Yeoman suffers from its extreme brevity (4 pages) so that it is little more than a list of central themes and a signposting to the various fields of research (covered in the preceding chapters) which are contributing to these themes. Some critical discussion of perspectives would have been of real value here.

Setting aside these minor blemishes the volume is timely, well conceived, and scholarly written. It fills a real gap in botanical literature. In general the chapters are uniform and appropriate in standard so that they will be of value to established scientists looking for a well documented survey but equally appropriate for advanced undergraduate reading. The authors and press are to be congratulated on the excellent plates and informative line diagrams many of which have been specially prepared for the volume.

H. E. Street

H. E. Street is Professor of Botany at the University of Leicester, UK.

Energy past and energy present

Man, Energy, Society. By Earl Cook. Pp. xi+478. (Freeman: San Francisco and Reading, May 1976.) Cloth £10.40; paper £4.50.

THE availability of energy is clearly a key element in the history of civilisations, the size, form and behaviour of societies and the choices open to them. Following in the steps of Fred Cottrell's remarkable *Energy and Society* (1955), the geologist and geographer Earl Cook here re-explores this fascinating set of connections. The result is a valuable, broad introduction to 'energy' in a historical, social, geographical and economic setting. The sweep is immense, ranging from the emergence of agriculture to future energy alternatives of varying sociopolitical stability. The approach is often refreshingly catholic: thus, a major section on the global distribution of energy resources starts with arable and pasture land; food (and its getting) commands as much attention as the major fossil fuels; the ethics of resource depletion and the sociology of control of energy use are discussed as thoroughly as the usual technical and economic topics. One consequence is to make the book a valuable source of inaccessible statistics and data, such as the efficiency of early machines; the total system energetics of transport, food production, mining and other industrial activities; and the rising energy and other costs of winning lower grade mineral and fuel resources.

Although useful up to a point, any uni-dimensional view of the world can all too easily be taken too far into empty truisms and gross overstatement. Cook all too often falls into these traps, walking dangerously close to a pure 'energy theory of value'. Statements such as "a doubling of living level requires approximately a doubling of available energy" (p190); that if energy use per capita is not increasing then economic development is reversing (p218); and that the gap between developed and underdeveloped regions in per capita energy use is widening (p18), underpin many of the book's theses, yet are demonstrably false. Nor is it exactly helpful, in the absence of any sustained discussion of overt political factors in the control of energy or of society, to come across remarks such as "high energy man seeks diversion in sports, business, drugs, sex, rioting, and crime" (p221), or that the Crusades could not have happened without the mobilisation of energy surpluses in Northern Europe "reflected in horses, armour and men"

(p195). Yet read with a wary eye and a few pinches of salt, this is nevertheless a valuable book as both a thought-provoker and source of information.

Energy Resources and Supply. By J. T. McMullan, R. Morgan and R. B. Murray. Pp. xii+508. (Wiley-Interscience: London and New York, March 1976.) £12.50; $27.50.

FEW of the 'broad survey' books on energy that are now flooding the market focus with such admirable clarity, thoroughness and wide scope as does this on the hard-nosed physics and engineering of the subject. Aimed at university courses and the more numerate specialists, the emphasis is predominantly on techniques and their physicochemical basis—in many ways a refreshing change that helps fill an important gap. There is also a strong bias towards the 'mainstream' supply and conversion technologies, with about half the book devoted to resources, extraction, processing and characteristics of the fossil fuels and the nuclear fuel cycle, both fission and fusion. Opening chapters survey the energetics of the biosphere-atmosphere system and photosynthesis, whereas there are relatively sketchy chapters on solar energy, other renewable sources, storage, energy-in-use and (with proper enthusiasm) the heat pump.

The coverage and density of information is impressive. Yet there are curious gaps. For example, there is much about peat as a large 'fossil' resource but nothing about the vast annually renewed fuel sources in (tropical) estuarine biomass; in a long section on nuclear reactors there is the scantiest reference to comparative efficiencies of fuel use; the account of solar collectors bristles with algebra but there is hardly a word about overall system design or integration of solar with other ideas (including storage and heat pumps), the keys to success in this area. More seriously, there is such a determined avoidance of almost everything but physics and engineering that the real world sometimes seems light years away. There is almost no mention even of costs, and the broader economic and social aspects of energy are confined to a single chapter of six pages. The total misunderstanding of the energy situation, needs and constraints of the undeveloped world, revealed in the final pages, is no less than appalling in a book which in many ways deserves to become a basic text for undergraduate and graduate students in the physical, engineering and environmental sciences. **Gerald Leach**

Gerald Leach is a Senior Fellow of the International Institute for Environment and Development, heading a program on alternative energy strategies for both the industrialised and less developed world.

Plant ecophysiology

Physiological Plant Ecology. By W. Larcher. Translated by M. A. Biederman-Thorson. Pp. xiv+252. (Springer: Berlin and New York, 1975.) DM 46; $18.90.

IN revising and translating into English the original German text *Okologie der Pflanzen*, the opportunity has been taken to adopt a more accurate title, which is supported in the Preface by an adequate description of the subject matter and aims of the book. Ecologists urgently seeking physiological information relevant to field situations for teaching and research purposes, invariably have to search a voluminous literature. Professor Larcher is to be congratulated on making this task so much lighter, by producing a concise account of many important aspects of plant ecophysiology.

Inevitably, the topics covered reflect the authors's own research interests and the limitations of space and cost. The approach adopted of portraying the development of the subject in the choice of illustrations and tables is ex-

tremely interesting. Much of the material is new and valuable syntheses have been made in a number of tables, yet in certain sections, especially those on nitrogen and mineral elements, the treatment is somewhat out-of-date. Under the heading 'Dry-matter Production' (p63) the section on growth analysis would greatly benefit by additional references to recent alternative articles and texts. There are very few obvious errors in the text: (cf p000) should be (cf p134) on p54 and (cf p000) following Id=direct solar radiation (p191) should be omitted from this revised translation. I would prefer to see Leguminosae used on pp94, 101, 115 and 119 instead of Fabaceae.

The book is a valuable source of reference to articles normally hidden away in the German literature and it is to be regretted that the current rates of exchange will put it beyond the pockets of most undergraduates.

K. Taylor

Dr Taylor is a lecturer in the Department of Botany and Microbiology at University College, London, UK

obituary

Jerome Vinograd was born February 9, 1913 in Milwaukee, Wisconsin, and died July 3, 1976. After undergraduate work at the University of Minnesota, he studied colloid chemistry with H. Freundlich from 1931 to 1935, first in Berlin then in London. He completed his doctoral work with J. W. McBain at Stanford from 1937 to 1939 and was associated with the Shell Development Co. from 1941 to 1951 where he worked mainly on problems of surface films and colloids. At that time, he gave up a secure industrial career and came to the California Institute of Technology because of his intense desire to work in the exciting new area of molecular biology. His early work at Caltech dealt mainly with the physical chemistry of proteins.

His first outstanding contribution to molecular biology was the invention, in collaboration with Meselson and Stahl, of the method of equilibrium density gradient centrifugation of nucleic acids in caesium chloride solutions. This method has been elaborated in Vinograd's and other laboratories to take advantage of the effects on buoyant density of base composition, denaturation, alkaline titration, binding of metal ions and of dyes, and hydration. All told, equilibrium buoyant banding has been one of the key techniques in the explosive development of the nucleic acid side of molecular biology since the mid 1950s.

Vinograd's outstanding work in the past 12 years has been in the study of closed circular DNA. The distinction between supercoiled closed circles and open circles was recognised in 1965. The ethidium bromide method of buoyant banding for the isolation of closed circular DNA was invented in 1967. This method, which makes it possible to analyse for and to isolate closed circular DNA in the presence of massive quantities of linear and open circular DNA, is widely used, and is one of the main reasons why there has been such exciting progress in the study of

closed circular DNAs in different kinds of cells.

Later work dealt with the properties of mitochondrial DNA from malignant cells and with the study of the replication of mitochondrial DNA, initiated by the discovery of D-loops in Vinograd's laboratory. More recently his students have made several important contributions to the studies of the properties of the nicking and closing enzyme which relaxes supercoiled DNA while leaving it covalently closed. The Vinograd group as well as the groups of Keller at Cold Spring Harbor and Wang at Berkeley observed that one can use gel electrophoresis and relaxation by an enzyme to count the number of turns in a closed circular DNA, and that there is a Boltzmann distribution in the number of superturns in molecules which are closed, under equilibrium conditions.

Vinograd was, by early training and in his approach to molecular biology, a physical chemist. He had a flair for recognising when an anomalous and unexplained observation, if subjected to fundamental physical chemical analysis, would lead to an unexpected and important result in molecular biology.

I have seen the special Vinograd touch many times. A simple example is the development of the velocity band centrifugation method by Vinograd, Bruner, Kent and Weigle in 1963. Shortley before, Jean Weigle had discovered that he could layer a dilute solution of viruses in aqueous buffer on to a sucrose solution (not a gradient) in a centrifuge tube, centrifuge and get a good sharp virus band. Vinograd was puzzled as to why the band was sharp, and reasoned that there must be some effect providing convective stability. He recognised that there had to be a self-generating (by diffusion) density gradient to provide convective stability; a careful study then led to the development of this general and widely used method.

I remember vividly a seminar at Caltech in late 1966 or early 1967 when J. B. LePecq described his studies on the binding of ethidium bromide (Etd Br) to DNA. An important point was that, unlike the common acridine dyes, there was a reasonable amount of binding at high salt concentration. Vinograd pointed out at the discussion which immediately followed that if the dye ion bound in 6 M CsCl it would cause a large shift in the buoyant density of the DNA because it would displace a Cs^+ ion. Experiments by W. R. Bauer (then a graduate student) showed that Etd Br would indeed decrease the buoyant density of DNA, but that the shift was much larger for a relaxed than for a closed circular DNA. On learning of this result, Vinograd immediately perceived the correct explanation in terms of the topological constraint on the unwinding of closed circular DNA, thus leading to the development of an extremely useful method for isolating closed circular DNA, as well as for studying the free energy supercoiling.

He suffered a major heart attack in 1954 and a second one in 1969. He accepted the resulting restrictions on his overall activities philosophically. But his condition did not affect the intensity with which he devoted himself to his scientific work. I remember him through all these years being in the laboratory regularly on evenings and weekends discussing and analysing experimental data in painstaking detail with his students, seeking a fundamental explanation for unexpected and unexplained results. His students and colleagues will miss his penetrating insight into technical problems and his wise counsel on general policy issues.

Vinograd was made a member of the National Academy of Sciences in 1968, and received numerous awards. He is survived by his wife Dorothy, and his two daughters Julie and Deborah by a previous marriage.

Norman Davidson

Person to Person

The Beilby Medal and Prize will be made to a British scientist for work in any field related to the special interests of Sir George Beilby (chemical engineering, fuel technology or metallurgy). Apply to The Convener of the Administrators, Sir George Beilby Memorial Fund, The Royal Institute of Chemistry, 30 Russell Square, London, WC1B 5DT.

The Commission of the European Communities is sponsoring travelling Fellowships for obstetricians, neonatologists and research workers who are actively engaged in perinatal monitoring. The successful applicants will spend three weeks in a centre for perinatal monitoring within the EEC, outside their own country. Applications to Commission of the European Communities, Directorate-General Research, Science and

Education, X11/C-1, 200, rue de la Loi, 1049 Brussels.

Address and details of the work of Rose Selavy on stellar lightning (ref. Brecher, K., *Nature*, **261**, 542; 1976) would be appreciated by Mr E. W. Crew, 26 St David's Drive, Broxbourne, Herts, UK.

There will be no charge for this service. Send items (not more than 60 words) to Martin Goldman at the London office. The section will include exchanges of accommodation, personal announcements and scientific queries. We reserve the right to decline material submitted. No commercial transactions.

nature

September 16, 1976

Still plenty for the BA to do

ANOTHER meeting of the British Association for the Advancement of Science (BA) comes and goes, and yet again the question arises, do we still need the BA, and if so, for what purpose?

Certainly, the annual meeting is not the only activity of the BA. Throughout the year there are study groups, science fairs, schools conferences, visits, publications, and so on, and it is indisputable that much of value is achieved by all this. But a lot of energy goes into the annual gathering, at which a good accumulation of speakers is assembled. To judge by the size of the audience at many of these talks, this effort is simply not being repaid by enough public interest to warrant an endless annual commitment. Is the fault with the BA and its organisers or is it elsewhere?

It would be easy to blame the BA. "Meet Britain's leading scientists" ran the headline on the poster put about locally in Lancaster. Directly beneath it was a large picture of the association's irrepressible and telegenic executive secretary, Magnus Pyke. Some of Britain's leading scientists, whose contributions are essential if the BA is to survive, might reasonably wonder whether the message should not have read, "Meet Britain's best known scientist".

The media do not help. Armed with copious preprints of lectures, correspondents find it all too easy to turn in large quantities of uncritical copy of the "said Dr Smith" type, which manages at once to convey the impression both of a lack of authority and of exaggerated claims. This conceivably does more harm than good in helping the public to a better appreciation of the achievements and limitations of science; it conceivably also scares away potential friends of the BA among the scientific community.

But it is an easy escape to blame the BA, its slightly amateurish image, its press coverage, for a major deficiency in us, the scientists. The BA still manages to persuade a good number to come and deliver lectures, but with many, no sooner is the lecture delivered than it is a case of up and away, probably to be back in the laboratory as soon as possible. Maybe there are pressing engagements elsewhere, but perhaps we have things to learn from a diverse audience every bit as interesting as they have to learn from us. And, of course, the number of practising scientists who go to BA meetings if not invited to speak is all too small. As a result, the annual gatherings seem to lack the necessary critical mass of those actually producing the subject matter. Such a cavalier attitude towards specialist conferences would, of course, be unthinkable.

It may be argued that scientists have better things to do than spend several days adding to their general scientific knowledge. If there was much evidence that scientists used other ways of preventing themselves from getting caught in a specialist rut, then that might be a perfectly valid argument. But we are dubious that the average research scientist is anything like as well read as he or she should be to take full advantage of all the intellectual opportunities that science has to offer. And we would not be surprised if this narrowness is significantly more pronounced in the United Kingdom than it is in many other parts of the world, no doubt largely as a result of an educational system which, for better or worse, produces single-minded, specialist experts at the age of 24.

There was a time, and not too long ago, when this was something to be proud of. But with changing economic circumstances people are beginning to ask whether, without too much lowering of the standards of scholarship, it might be possible to ensure that the scientist's interests are not quite so intensely focussed. The BA should undoubtedly have a role to play in this by providing an arena in which scientists can learn from each other, from administrators, from politicians and even from taxpayers. This requires a broader commitment, of a kind which there once was, but which is now lacking. Is there any chance that Britain's scientific community might regain it in the future? □

National Air and Space Museum, Washington, DC

America's history lesson

Chris Sherwell sees the Smithsonian's latest spectacle in Washington, the National Air and Space Museum

TO glide from Washington's white heat of late summer, past the imposing pink marble and tinted window facade of the National Air and Space Museum, and on into the air-conditioned splendour of an enormous hall dedicated to the milestones of aviation, is something of an experience. For there, in just one room, is the Wright Brothers' Kitty Hawk; Lindbergh's Spirit of St. Louis; the first plane to break the sound barrier; and the X-15, which has flown at Mach 6.72. And that is just on the 'Air' side. For 'Space', and in the same room, there is a replica of Sputnik and a duplicate of Explorer I; John Glenn's Mercury capsule, Friendship 7; Gemini 4, which carried the first US space walker; a replica of the Mariner 2 interplanetary probe to Venus; the Apollo 11 command module, Columbia; and the first piece of "touchable" lunar rock. It is all a remarkable introduction to a remarkable display.

Nearly three million people have already traversed that hall, and the museum opened only eleven weeks ago; on one particular day, 82,000 people decided to visit. All seek to be informed and educated in an interesting and entertaining way about the history of aerospace. Most leave satiated. The man who piloted 'Columbia' on that memorable moon landing flight, Michael Collins, is the museum's director. He acknowledges that there is no accepted view of what a museum is supposed to do, but at his disposal is a team of 250, a vast archive, a huge 'spacearium', an enormous theatre with a 55-foot high screen, scores of audio-visual aids and, of course, hundreds of artefacts—which means aircraft and spacecraft. The result is a 23-section history of flight, beginning with balloons, and ending with the Universe, with digressions into flight and the arts, air-traffic control and the benefits of flight on the way.

Although the balance of the museum's overall display is in favour of Air rather than Space, the most spectacular exhibits are the space-related ones. The Skylab Orbital Workshop and Multiple Docking Adaptor, together with the giant Minuteman and other missiles, dominate the second major hall, the Space Hall. Next to these are the Apollo–Soyuz spacecraft and docking module; nearby is a Lunar Module, with the Orbiter and Surveyor satellites suspended above. A room dedicated to the Moon landing project includes a complete rundown on all the crews of manned US spacecraft, a full-size mock-up of the Lunar Module cockpit, displays of experiments performed on the Moon (along with the Lunar Roving Vehicle) and several samples of lunar rock.

None of this could have happened without rockets like the Saturn V, at 353 feet so high that there is room in the museum for only one of its main engines; mirrors have to be deployed to give some idea of the overall size. Rooms on rocketry and rocket history accord due prominence to Goddard, Oberth and Tsiolkovsky, but the television series *Star Trek* is not forgotten either. The superb satellite room has communications, Earth monitoring, science and weather satellites, but no spy satellites. Collins would like one, but no one acknowledges they exist, and no one has recovered one.

Air sections too have their own attractions and distractions. Apart from the World War rooms, there is, for the more ghoulish, a constantly running film of the last flight of the Hindenburg and film of the more spectacular of the crazy acrobats who flew on planes rather than in them. In the Sea-Air Operations rooms visitors can stand on the bridge of an aircraft carrier and watch planes land on the deck, or alternatively (and more alarmingly, since there is no attempt to explain the military concept behind an aircraft carrier) join the pilots of the planes themselves as they go on a short 'mission'. Elsewhere, certain technical principles are exceedingly well demonstrated. There are explanations of the propeller, the turbofan engine, jet propulsion and, in the Flight Technology room, of supersonic flight and even of flight itself.

As always, the problem, even in an aerospace museum, is space—that is, the lack of it. The museum is, in fact, only two-thirds of its originally proposed size. Thus, there has to be some disappointment for everyone. In the third major hall, the Air Transportation Hall, for example, a Douglas DC-3 hangs majestically, "the most important single aircraft type in the history of air transportation". The same could not be done for what may turn out to be the most important planes since—the Lockheed C-5A and the Boeing 707. The problem, of course, is that there is no room for what would need to be another airport in the centre of Washington.

How, then, did Collins and his team decide priorities? "It's an imperfect process", he says; some categories—First and Second World War aviation, for example—chose themselves. One problem, revealed by an early survey of audiences in four of the halls, is that people by-pass important exhibits in their desire to see particular attractions like the Spirit of St. Louis. That exhibit is already the most popular, and Collins suggests that this is something to do with Lindbergh himself: people can relate to his "one-man show" a lot better than they can to the space programme, where there is "a cast of hundreds of thousands".

In fact, although the enormity of the US space effort is more than adequately conveyed by the sheer size of some of the exhibits, the impact which the aerospace industry, indeed flight itself, has had on American people and Western society generally is perhaps insufficiently emphasised. Partly, this is a problem facing any museum, where the idea of history as events, turning-points and personalities finds its greatest expression. But it is also a matter of perspective.

Take one particular personality, Howard Hughes—surely the *eminence grise* of US political and economic life. On the tiny Giants of Air Transportation exhibit the briefest of blurbs says only: "Howard Hughes for many years was the guiding force behind TWA and worked on the design of the Lockheed Constellation". In the larger Flight Technology room, the Hughes H-1 Racer is the centrepiece, and is represented, proudly, as "a major milestone on the road" to Second World War fighters like the the Grumman Hellcat, the Japanese Zero and the German Focke-Wulf FW 190. And in the Benefits of Flight room, Hughes Aircraft and the other major aerospace corporations like Boeing, Lockheed, Rockwell, Northrop, Grumman and McDonnell Douglas, merit just a mere mention on a backdrop list. Nowhere in the museum can the links, if any, between war, politics, the economy and the major corporations be more than guessed at or presumed.

Crucial decisions of the present—in the case of the US, those concerning the so-called "Sale of the Century" of fighter planes, the B-1 Bomber or the Boeing 7X7 and 7N7—receive little emphasis. There are some elements of topicality—a full-scale model of the Viking Lander, for example. But the biggest gap occurs with the supersonic transport, or SST, of which there is not even a passable model. The decision not to go ahead with this in the USA was, in the view of many, a defeat for science and technology. Not to have an exhibit is a defeat for a museum which aims to look to the future as well as the past, and which actually manages this, if somewhat disappointingly, with the Space Shuttle. Collins himself admits that this may be a weakness, and even allows that, as far as topicality is concerned, the museum "does not do that very well".

Not that the museum fails totally to overcome the problem of presenting history in terms of events. But it does go only a little way towards describing the social impact flight has had. The Benefits of Flight rom is an acknowledgement that the major exhibits do not offer aerospace automatic self-justification. There the view is expressed—near a deactivated atomic bomb—that flight has contributed to peace rather than war, at least between the major powers ("the sword over Damocles' head"). Exhibits in the same room refer to work on the Earth's resources, to scientific research, to agriculture, to the weather, to communications and so on. A juke-box blaring out Elton John's "Rocket Man" and Frank Sinatra's "Come Fly with Me" spotlights the cultural impact; so does the list of additions to our vocabulary, and the whole Arts room.

But in the areas where the greatest human transformation has been wrought—in population mobility, in the way people think of time and space and, above all, in the way production is organised—the museum shows its greatest weaknesses. How to remedy this? Perhaps the answer lies in the spacearium and theatre, which currently come closest to surmounting the difficulty with the presentations "Cosmic Awakening" and "To Fly", both pitched to a mass audience, and a mass American audience at that.

Above the two floors of display, together with the staff offices, is the library, available by appointment for any researcher to use. It has provided the main documentation for the show below. Collins says he has sought neither to editorialise nor to impose too much uniformity through excessive analysis. He favours an element of surprise instead, treating some subjects more seriously and some more whimsically than others. He agrees that the aim must be to make a presentation that allows an audience to draw its own conclusions, rather than one which makes particular preconceptions and makes certain conclusions inescapable. He acknowledges that some things need to be dressed up just to catch attention, but would hope that the inevitable built-in bias of any museum is in this case acceptable.

Apparently, it is: if the anti-space or anti-business lobbies in the United States believe that he has in fact failed in his task of educating a curious public, their views have gone largely unrecorded. Collins and his staff are pleased with the press they have received. The only serious regret about such a venture must be that it is a National Air and Space Museum, and not an international one. □
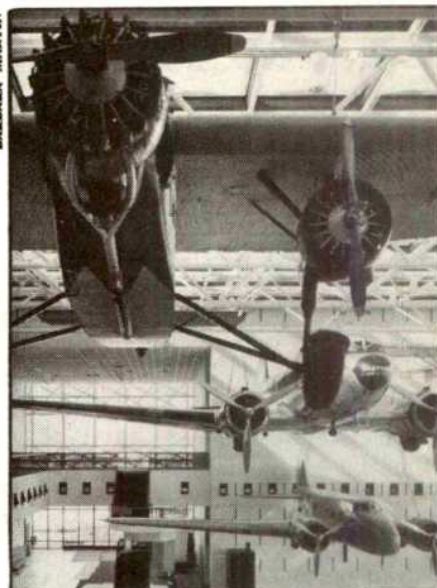
Hall of Air Transportation
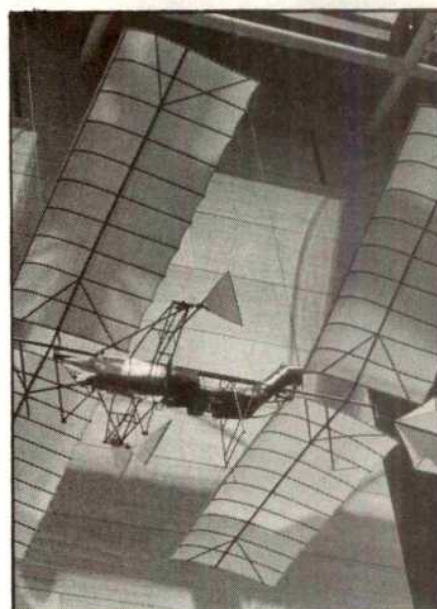
Steam driven aerodrome
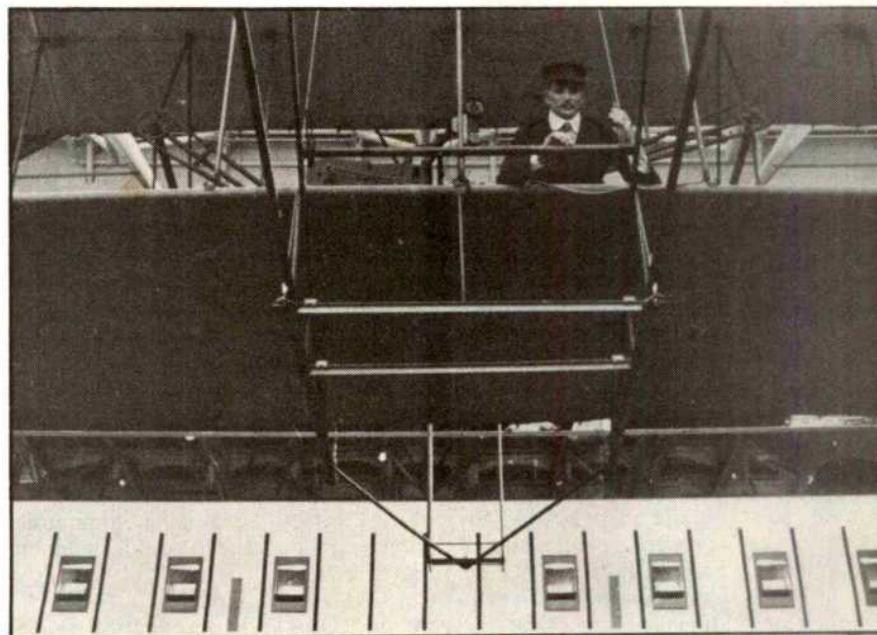
The Wright Brothers' original Kitty Hawk Flyer

ESA airs itself

Last week, the European Space Agency displayed itself at the Farnborough
Air Show in Britain. **Allan Piper** reports

IN solid commercial terms, this year's Farnborough International Air Show brought relatively little joy to the world's aerospace industries, with firm orders for hardware considerably down on recent years. To the many large corporations affected—companies such as Rolls-Royce and Westland from Britain, Aerospatiale and VFW-Fokker from Europe, and Boeing, Lockheed and McDonnell Douglas from the US— the cutback is no less unwelcome for all that it was expected and is seen as only temporary. But few of the contractors present will regard the latest Farnborough as a complete failure. In the first place, the long-term future of the aerospace industry seems secure enough, and this year's record crowds again proved the show's worth as a massive public relations exercise. Second, and far more important, the get-together encouraged a measure of serious discussion on what has become a vital issue for world aviation—international collaboration.

With the enormous expense now involved in the design and development of most civil and military aviation hardware, it is scarcely surprising that collaboration emerged very early during this year's show as the new watchword for aerospace. In Britain, attention focused particularly sharply on planned Anglo–US discussions announced by Gerald Kaufman, the UK Industry Minister.

But for the 12-nation European Space Agency (ESA), prominently represented at Farnborough, collaboration is an already well-established principle. Since its launching almost 16 months ago ESA has put no less than eight scientific satellites successfully into orbit, working with three giant consortia involving many of Europe's aerospace contractors.

While that prolific record itself speaks volumes for the success of European cooperation, early discussions on joint ESA–NASA ventures have ensured that broader collaboration will not go by the board either. The Spacelab project, already under way, represents ESA's largest scientific contribution to the NASA space shuttle programme, planned for the late 1970s and 1980s. Next month, ESA officials are expected to approve funds for another joint project with NASA, the Large Space Telescope. And many of the other new projects in the ESA pipeline will play important roles in forthcoming international programmes.

Meanwhile, Farnborough provided

ESA with an ideal opportunity to stamp its presence on the public's mind. Operating comfortably on a current annual budget of $600 million, the agency has much to offer in the way of eyecatching space technology. Triumphs have already been notched up with the ESRO, HEOS and COS B satellite programmes, and ESA has an equally impressive timetable drawn up for the next four years or so. New ESA projects include:

● OTS, the Orbital Test Satellite, scheduled for launching in mid-1977. Primarily a demonstration model for the European Communications Satellites (ECS) programme, OTS is nonetheless adaptable for various missions. The basic service module can, for example, carry weather forecasting equipment just as effectively as a telecommunications payload. The first operational ECS's will provide around 5,000 telephone circuits and two wideband television channels along with data transmission and teleconferencing services. During Farnborough a £900,000 design contract for ECS-1 was placed with MESH, one of the European consortia.

● METEOSAT, a long range weather forecasting satellite, scheduled for launching next year. METEOSAT will provide Europe's contribution to the Global Atmospheric Research Programme and the World Meteorological Organisation's later World Weather Watch. It will also fit into a global coverage involving Japan, the Soviet Union and the US.

● GEOS, the world's first scientific geostationary satellite, will monitor the magnetosphere. Scheduled for launching next year, it will play a major role in the International Magnetospheric Programme.

● IUE, the International Ultraviolet Explorer, a joint venture between NASA, the UK Science Research Council, and ESA, to be launched next year. ESA will contribute the solar-cell array for the satellite, and will design, build and operate the ground station near Madrid in Spain.

● ISEE-B, the International Sun–Earth Explorer, part of another NASA–ESA collaborative project. ESA's ISEE-B will be launched in tandem with NASA's ISEE-A next autumn.

● MAROTS, a shipping communications satellite scheduled for launching early in 1978. The payload will sit on an OTS service module.

● AEROSAT, an aeronautical satellite for ground to air communications on
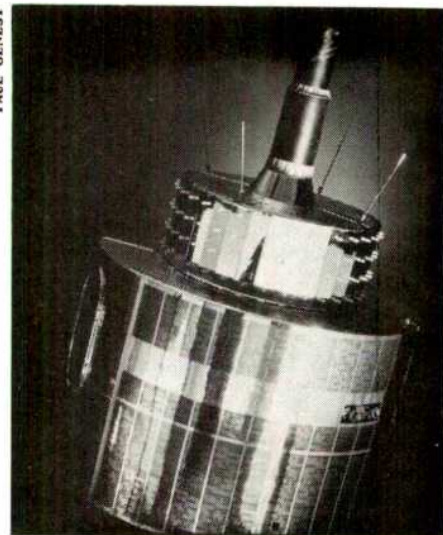
Earth. Two AEROSATs will be launched into geostationary orbit at the end of 1979 and beginning of 1980.

● EXOSAT, an X-ray astronomy satellite that will monitor cosmic X-ray sources. Along with GEOS, IUE and ISEE-B, EXOSAT will be a purely scientific spacecraft. It will be launched in 1980.

● Spacelab, the star of them all. A re-usable pressurised space laboratory, Spacelab will house scientists and engineers working in shirtsleeve comfort. A major component of NASA's space shuttle programme, it will become operational during the 1980s.

● Increasing congestion in the international queue for NASA's launching facilities, which is likely to become worse as NASA steps up its own requirements, has led ESA to develop a heavy launcher, Ariane, for itself. Designed to lift satellites weighing around 800 kg, it could prove crucial to the ECS programme. Ariane's four qualification shots, planned for 1979 and 1980, will be from the permanent launching base in French Guiana. □

PAUL GENEST

METEOSAT

ESA

GEOS

USA

Approaching scientific issues

While the Ford administration has been setting up the White House Office of Science and Technology Policy (see page 184), Jimmy Carter's campaign organisation has been establishing links with prominent members of the scientific community. Colin Norman reports

THE campaign between Gerald Ford and Jimmy Carter for President of the United States officially kicked off last week with a round of speeches from the two candidates and their Vice-Presidential choices. Nobody said anything about science, but then nobody really expects scientific issues to play much of a role in this campaign. "Science", as one adviser to the Carter campaign pointed out, "is not exactly the most burning issue in the public mind". Moreover, there's little discernible difference between Republican and Democratic policies for research and development.

But that doesn't mean scientists will play no part in the campaign. Far from it. In fact, it has become something of a tradition for large number of scientists and engineers to shed their usual non-partisan roles and enter the fray on behalf of one candidate or another, and this year will be no different. An extra facet in this campaign, however, is that the Carter organisation has been actively seeking advice from a diverse array of scientists, and that advice has been flowing in quite freely.

It is coming from a very loosely knit science policy task force, whose activities are coordinated by Lewis Branscomb, chief scientist for IBM, former director of the National Bureau of Standards and a well respected figure in science policy circles. Other members of the task force include such equally prominent figures as Harold Brown, President of California Institute of Technology, Harvey Brooks, Professor of Technology and Public Policy at Harvard, David Baltimore, a Nobel Prizewinner from MIT, and George Low, President of Rensselaer Polytechnic Institute.

The science policy task force mirrors similar units advising Carter on a vast range of other issues. It is a very informal, *ad hoc* arrangement involving members who rarely, if ever, meet. Instead, they draft policy papers which are sent to the Carter campaign headquarters, to be used as Carter's staff sees fit in drawing up position statements and speeches. As one task force member noted last week, "we haven't exactly got much power in formulating

the policies". This conforms with Carter's well-established approach of seeking opinions from a diverse array of individuals before enunciating his own positions. In foreign policy matters, for example, his advisers include people whose perspectives differ so widely that they must be offering all sorts of conflicting advice which the Carter campaign organisers are filtering and stitching together.

Branscomb said last week that he first became acquainted with Carter late last year, and that he was "impressed with (Carter's) intelligence and approach to scientific issues". Branscomb said that he personally favoured Carter's nomination as the Democratic candidate, and agreed to head the science policy task force when Carter approached him. A measure of the generally non-partisan nature of science policy here is that Branscomb was also a member of a committee advising the Ford administration on an agenda for the new White House Office of Science and Technology Policy (OSTP).

Branscomb has repeatedly emphasised the informal nature of the task force, and its broad membership. Many people, he said, are simply being asked for their views on issues, and their involvement carries no implication that they would necessarily endorse Carter's candidacy. "We are trying to make a contribution to good government," he told *Nature,* and added that the task force is "trying to avoid being an elitist, exclusive club with an inside track".

With Carter already commanding a formidable lead in the public opinion polls, it's clear that the task force is looking beyond November, attempting to lay the groundwork for the first few months of a Carter Administration. In that regard, there has been some speculation that if Carter is elected he will ask Branscomb to be his science adviser and the Director of OSTP. Asked last week whether he would accept such an offer if it is made, Branscomb said the possibility has never been discussed with the Carter organisation, and that he prefers to cross that bridge if he comes to it.

In spite of all the advice flowing into the Carter camp on science policy, so far there is little major difference between the two parties on scientic matters. As far as the party platforms are concerned, for example, though the Republicans devote a few paragraphs to science and technology, the wording is not at all controversial, and could easily have been part of the Democratic manifesto.
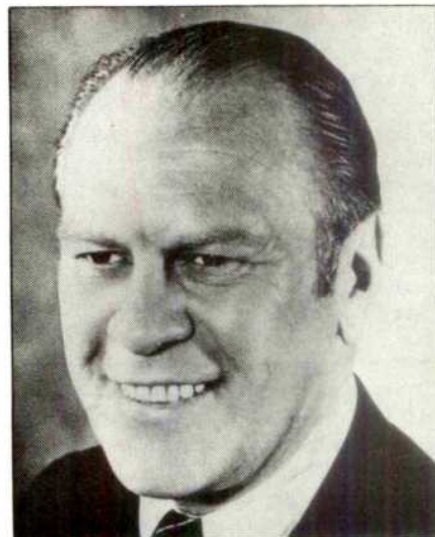
Ford, disastrous

Carter, looking forward

The Republican document states, for example, that "every aspect of our domestic economy and well-being, our international competitive position, and national security is related to our past and present leadership in basic and applied research and the development of our technology. But there can be no complacency about our continued commitment to maintain this leadership position". The Democratic platform simply makes no reference to such matters.

Some differences are, however, beginning to emerge between the two candidates on energy and environmental matters, which have some bearing on research and development.

In a speech before the National Press Club last year, for example, Carter attacked "President Ford's disastrous energy policy", which he described as the "Ford/oil industry energy policy", and claimed that it would eventually lead to higher energy

prices and greater reliance on oil imports. He said he was looking forward to the opportunity to formulate his own energy policy, in an atmosphere devoid of crisis. Though he hasn't been too specific about his proposals, Carter has since suggested several times that he wants to see more emphasis on conservation, more spending on solar energy, and more reliance on coal. He has also attacked the Ford Administration's plans to push development of a synthetic fuels industry.

The chief difference between the Carter and Ford energy policies lies in the area of nuclear power. In a statement made to a United Nations conference last May, Carter suggested that dependence on nuclear power should be held to the minimum necessary, and he has said that he would reassess the priority being given to the liquid metal fast breeder reactor—the most expensive single energy research and development project supported by the Ford administration. He has, however, stopped short of calling for a moratorium on nuclear power, and he refused to support the California initiative calling for a virtual halt in the development of nuclear power there. (The initiative was defeated in a state-wide referendum last June.)

As election day approaches, science and technology are unlikely to figure very prominently, simply because they do not figure very prominently among the issues about which people are concerned.

If that is a source of encouragement for scientists, it is largely because of the prestige they enjoy. □

AFTER a series of unexpected and often bizarre delays, the newly-re-created White House Office of Science and Technology Policy (OSTP) is finally getting down to business, four months after it was legislated into existence and—if the opinion polls are to be believed—only a few weeks before the Ford administration is scheduled to make its political exit. Even though the administration's days may be numbered in double figures, however, OSTP is being established as if its working arrangements will endure well beyond next January's Presidential inauguration.

H. Guyford Stever, who heads the office and who also holds the title of Science Adviser to the President, said in an interview last week, "I'm not expecting a change of administration. And neither is the President". Accordingly, he is busy recruiting staff, hiring consultants and establishing committees, which means that Jimmy Carter, if he is elected, will inherit a functioning science policy office and some top-level committees whose members would be difficult to replace.

Although legislation establishing the office was signed into law by President Ford on May 11, after taking more than a year to pass through Congress, the OSTP's debut was delayed by an unexpected dispute over the appointment of a director. It had long been rumoured that Ford would nominate Simon Ramo, head of the giant aerospace and defence consortium TRW, as his science adviser and director of OSTP. But the nomination was never made, apparently because Senator Kennedy's office insisted that Ramo would have to divest his holdings in TRW before the Senate would confirm his appointment. Ramo decided that, since the polls all pointed to a resounding Ford defeat in November, it wasn't worth the price to be a science adviser for only a few months. Attention then turned to Guy Stever.

An aerospace engineer and former President of Carnegie-Mellon University, Stever has been Director of the National Science Foundation (NSF) since 1972 and part-time White House Science Adviser since mid-1973, when Mr Nixon scrapped the old Office of Science and Technology and consigned some of its functions to the NSF. Stever would therefore seem a logical choice for the full-time post of Presidential Science Adviser. But, when news of his impending nomination filtered out, four right-wing Republican Senators warned Ford that they would oppose the appointment. Led by Senator Jesse

Man at the top

H. Guyford Stever

Helms, they issued a statement criticising Stever's stewardship of the NSF, resurrecting many of the disputes which have swirled around the NSF's education programmes for the past year.

The opposition surfaced while Ford was locked in his neck-and-neck with Ronald Reagan for the Republican nomination. Stever, mindful of the fact that Ford was trying to steer clear of fights with the party's right wing, suggested that his nomination be reconsidered and the matter was put on the back burner for nine weeks. Then, when Ford had the nomination virtually locked up, Stever's nomination was sent to the Senate early in August and swiftly approved with only token opposition.

Stever said last week that he agreed to take on the assignment with the understanding that he wouldn't remain long after January, no matter which administration takes office.

The legislation establishing OSTP also sets up a top-level advisory committee, consisting of between 8 and 14 people drawn from a variety of fields, which will conduct a two-year investigation of federal science support and priorities. One of Stever's first acts was to recommend that Ramo be appointed chairman of that committee, a suggestion which Ford promptly accepted. The other members will be appointed later this month. (Ramo's appointment, incidentally, will present an interesting situation if Carter becomes President, for Ramo was a co-chairman in 1972 of a committee of scientists and engineers supporting the candidacy of Richard Nixon.)

Stever has begun to staff the office with people drawn from NSF's science policy office and a few individuals on secondment from other government agencies. In addition, he has appointed two senior policy consultants, William Nierenberg, Director of the Scripps Institution of Oceanography, and Donald Kennedy, Chairman of the Program in Human Biology at Stanford University. Both will devote half their time to OSTP, assisting in the development of policy studies on biological and environmental issues. Other consultants, Stever indicated, will be called upon as needed.

Asked last week what he considers as OSTP's chief goal, Stever said, "some scientists think that the goal should be to increase the support for science, but they are mistaken. The important thing is to establish an easy relationship at the top level of our government for the scientific community to have an input" into national policy.   **Colin Norman**

USSR

● The return of the Salyut 5 crew without having set a new endurance record has aroused speculation abroad as to whether a premature return was necessitated by the psychological stresses of the voyage. This is doubtless due to the fact that it is the first time that the Soviet press has been so explicit about the monitoring of the cosmonauts, which, it appears, involved the analysis of all routine communications with ground control to detect signs of stress.

Cosmonauts Volynov and Zhlobov wanted to be kept up-to-date with world news and sporting events as well as enjoy the on-board entertainments of taped music and a set of photographic slides. In the early part of the flight it was noted that "every detail about the progress of the Olympic Games gave them additional reserves of energy", and during the flight certain innovations were introduced—routine transmission from base were given a background of soft music and special emphasis was placed on "warmth and lively human interest in the concerns of the crew" shown by the communicators.

The official reports, however, do not suggest that the stresses developed were inadmissibly great; the new measures suggest an experiment in remote-control psychotherapy rather than a psychological emergency. Soviet space psychologists have long been concerned with the problems of in-flight sensory deprivation (including the lack of wide human contacts). Over the years this has led to a number of diverse selection procedures for would-be cosmonauts, ranging from the idea that candidates who report monochrome dreams are preferable to those who dream in colour, to the recent pronouncement that no more women will be accepted as cosmonaut-pilots, although they will still be eligible to serve aboard space-craft as "experts"—doctors, astronomers or stewardesses.

● A recent decree of the Central Committee of the Communist Party and the Council of Ministers of the USSR focused on "urgent measures for ensuring for the national economy fuel, electrical and thermal power, and the economic use of fuel and energy reserves". Addressed to all strata of society from the Councils of Ministers of the Union Republics down to the managers of individual enterprises, it called not only for all possible savings but also for the increased use of "peat, shale, briquettes and timber". It was somewhat reminiscent of a speech by Academician Vladimir A. Kirillin, Chairman of the State Committee for Science and Technology, during his visit to Britain in 1974. This extolled Soviet advances in the technology of fast breeder nuclear reactors, and also stressed that in the immediate future the Soviet power industry envisaged a considerable expansion of open-cast mining for low-grade coal.

The new emphasis on saving and on the use of low-grade fuels may be associated with recent hints of setbacks in the nuclear power plans. Writing in *Pravda,* Academician Nikolai A. Dollezhal (the chief

designer of the reactor for the first Soviet nuclear power station) indicates that there are still considerable engineering problems involved in the construction of nuclear stations, that the manpower involved in installing the equipment is some 2-3 times greater than for thermal stations, and that the efficiency of nuclear stations is generally less than that for thermal.

Moreover, he says, "there are grounds for assuming that in the not too distant future we shall be close to exhausting the 'ecological capacity' of the region of present siting of nuclear power stations." The new economy drive is aimed at ensuring the "unconditional" fulfilment of the five year plan; consequently neither industrial production nor fuel exports will be affected by the measures.

● '*Melioratsiya*', the ten-year-old land-improvement scheme involving drainage and irrigation projects and the general management of soil resources, has received a fresh emphasis. Following the disclosure of new plans over the past few months, *Pravda* has identified the scheme as "a most important task of the whole party, the whole Soviet nation".

This latest move to boost agricultural production includes the resurrection of far-reaching plans to divert northward flowing Siberian rivers south into central Asia and Kazakhstan. In the shorter term, the new measures include calls for substantial increases in the yields of cereal and other crops from "improved lands".

The renewed importance accorded to *melioratsiya* places involved workers high in the industrial hierarchy: incentives range from a special "feast day" for all concerned, to the title "Master of Irrigation" (First or Second Class) for outstanding individual contributions.

During the recent *Biosfera-76* exhibition in Minsk it was revealed more than 2 million hectares of marshland have already been drained in the Byelorussian SSR; and the past five years have seen an increase in the crop-yield of the reclaimed lands. Nevertheless, erosion is causing losses of up to 20% in crop production, and it is now proposed that some 75% of the peaty soils should be put under grass, leaving only 25% for cereal crops. With the continuing loss of productive soil to industry, and the extensive pollution of water resources also causing serious concern, the indications are that the new plans for *melioratsiya* will require considerable efforts if targets are to be met.

● The Soviet Union has agreed to cease whaling within the next two years, according to a recent statement attributed to Mr Nikolai Makarov, Chargé d'Affaires at the Soviet Embassy in Ottawa. This is a reversal of a long-standing Soviet policy, which maintained that a whaling moratorium would be positively harmful, since it would put an end to research on the biology of whales.

According to a spokesman of the USSR Ministry of Fisheries shortly before the 1974 Session of the International Whaling Commission, the established seasonal quota is insignificant, and there is "no biological necessity" to stop whaling completely since the USSR (which, together with Japan, takes 90% of the world's catch) "consistently and accurately observes all the provisions of the Whaling Convention, while attaching specially great importance to the development and implementation of effective measures to protect the whales".

At the 1976 session in June, the Soviet Union strenuously opposed the proposal to set quotas by "yield by weight", rather than by the "maximum sustainable yield" introduced in 1961. The USSR disapproved of the change, to be applied in the first instance to sperm whales only, arguing that the Commission had already done much in the past two years to maintain sperm whale stocks.

The reversal of policy is more than a cause for satisfaction among conservationists; hitherto, it is said, sperm whale oil has played an important role in the manufacture of Soviet tanks and missiles.   **Vera Rich**

IN BRIEF

Williams the DES

IN a Cabinet reshuffle Shirley Williams last week took over the post of UK Secretary of State for Education and Science from Fred Mulley. While Mr. Mulley's low-profile tenure of the post —marked by a period of economic stringency, cuts and out-of-work teachers—led to little in the way of long-term science policy, hopes are high among scientists and policymakers that Mrs. Williams will provide a different style of leadership, taking science under her own wing (rather than leaving it to a junior Minister) and generally infusing some optimism into what is at present a gloomy scene.

Mrs. Williams was a junior Minister for Education and Science between 1967 and 1969; she returns to the Department from her position as Secretary of State for Prices and Consumer Protection, held for the past two years.

More for US research

Government spending on science and technology in the United States is expected to reach $23,500 million during the fiscal year beginning October 1, according to figures recently published by the National Science Foundation (NSF). That would represent a modest increase from this year's level of support, even when inflation is taken into account.

The welcome increase follows a long period of decline in the federal science budget, as inflation has gradually eroded its purchasing power. The increase will be sufficient only to restore science support to its 1972 level, however, which in turn was 20% below that of 1967, the peak year of federal science support. The growth next year is almost solely attributable to large increases in funding for energy and defence research and development.

New orchid for Britain

Whatever its effects on wildlife in general, this year's warm dry spring brought a notable addition to the British flora: the orchid *Ophrys bertolonii*. A single specimen has been found near Swanage, Dorset, growing in undisturbed downland turf; how it got there will remain a mystery, unless some enthusiast confesses to having planted it there as a seedling some years ago. One of the loveliest and most unmistakable of European orchids *O. bertolonii* is essentially a plant of the western Mediterranean, although it occurs north of the Appenines and is also recorded from Bulgaria.

The notion that the British specimen has been in position for some years without flowering is strengthened by the discovery in Gloucestershire, also this season, of a second species, *O. sphegodes*.

IT must have come as a surprise to most people in Britain to learn that, according to the statistics for food consumption recently published by the UK Ministry of Agriculture, Fisheries and Food (MAFF), the nation's population is on average eating less than the minimum diet recommended by the United Nation's Food and Agriculture Organisation (FAO). Britons are apparently saved from starvation only by the efforts of the confectionery and brewing industries, whose products contain sugar that makes up for what would otherwise be a significant deficiency in calories.

It is difficult to take these figures seriously. If the average level of consumption is so near the starvation line, then something like half of the UK's population should show symptons of undernourishment, for the least-privileged members of society probably eat fewest sweets and drink least beer. Yet it is rare to see anyone except a criminally neglected child or an uncared-for elderly recluse, who is seriously underweight. One cannot walk along a city street in Britain without seeing many who are grossly overweight: The school medical service complains about the increase in the number of obese childen. Surveys show that as many as 75% of middle-class and middle-aged women, and over 50% of men, worry because they are overweight. Food advertisements push products that are low, not high, in calories. Newspapers are full of diets devised to allow consumers to lose weight without suffering the pangs of hunger. As a nation Britain shows all the signs of overeating, not of consuming

too little. The MAFF not only suggest that Britons eat too little food of any kind, but also that "the quality of food consumed shows a steady decline". This view is based mainly on data showing falls in the consumption of meat and other animal products, and of sugar. But as Dr. J. V. Durnin,

MAFF gaff

KENNETH MELLANBY

---

a nutritionist at Glasgow University, has commented in *The Times* newspaper, these changes should be viewed as a limited improvement, and not a deterioration in feeding habits. Moreover, they accord with a recommendation from the Royal College of Physicians.

The MAFF seems intent on perpetuating the fallacy that, because some protein is necessary for health, the consumption of more animal

products leads to better nourishment. Some time ago, the Minister even suggested that any reduction in the amount of meat eaten in Britain would give rise to serious malnutrition. Britain's diet may indeed be changing. Rising costs may be preventing everyone from consuming all the things they might wish to eat. Some people may be getting less pleasure from eating, or, more likely, from overeating. But in presenting dietary data, the MAFF should surely pay heed to the principles of nutrition, rather than to those of hedonism.

Those living in Britain can observe the real situation, but what will observers in other countries, or in organisations like the FAO, make of the MAFF statistics? Will the UK's 55 million citizens be added to the thousands of millions in countries even poorer than Britain already listed as suffering from food shortages? In the past, many statistics published by international bodies have exaggerated world starvation and malnutrition. It is revealing to find that equally exaggerated data can be produced for a country like Britain, where accurate information might be expected.

Yet there are too many hungry and even starving people in the world today. There are many countries where overeating is not, as it is in Britain, the main cause of 'malnutrition'. The problem can only be solved if there are accurate statistics for food consumption, and proper estimates of food requirements, for all countries. A start should be made by producing such information for Britain.

correspondence

Ball lightning

SIR,—Ball lightning is a relatively rare phenomenon, usually occurring under conditions of very strong electrical fields and usually associated with storms, though not necessarily in the storm area. The balls can be from one inch to several feet in diameter and can have various colours. Some explode violently and some fizzle out quietly. The balls move through the air, apparently independent of wind and gravity, and are known to enter structures—even through a metal window screen and into a metal aircraft. There is no good theory to explain ball lightning, but J. R. Powell and D. Finkelstein are working on a d.c. field theory in which a strong electric field adds energy to the ball to sustain it (*Structure of Ball Lightning*, Brookhaven National Laboratory, Upton, New York).

Little is known about the ignition hazard of ball lightning, but in the course of investigating accidents in the past ten years, we have heard the following statements from witnesses:
● Distillate was being loaded into the open top of an 8,000 gallon transport truck. "A ball of light travelled along the fill pipe and down the droptube. It entered the truck compartment and the truck blew up."
● Solvent was being loaded through closed piping into the bottom of a 7,600 gallon transport truck. "A ball of light followed along the fill pipe toward the truck and disappeared. Then the truck blew up."
● An observer on a ship several miles from an empty VLCC (very large crude carrier) tanker. "A ball of light travelled along the deck of the ship and disappeared. Then the ship blew up."
● Jet fuel was being loaded into the open top of a transport truck. "A ball of light entered the truck compartment and it blew up."
● A retired New York City fireboat captain, Charles Wilson of Codan Marine Inc., New York. "I know of three cases where a ball of light was reported to have moved across the deck of a barge just before the barge blew up."
● A tanker captain with the same corporation, Captain Wykoff, reportedly "saw a ball of light move across the deck toward a compartment before it blew up."
Taken individually, these are each incredible and are usually ignored by safety experts. Since there are these eight cases reported by separate observers, it appears that ball lightning might be an infrequent and unrecognised source of ignition.

There probably are other incidents, and we are soliciting any information that others may have. Our concern is that this is an unrecognised source of ignition in fuel transport operations and that conventional protective methods may be inadequate.

I. GINSBURGH
W. L. BULKLEY

Amoco Oil Research Department, Naperville, Illinois 60540

Forecasting error

SIR,—Kenneth Mellanby, in a short but cogent article ("Forecasting Error", August 5, page 441), is highly critical of the state of the art of forecasting many of the social and economic components of present society.

Most of his strictures are only too true. I would however take issue with his implication that "only some form of rigid totalitarianism could improve the statistics". This can surely be tested by examining the forecasting skills of societies which have for some time displayed this characteristic. To my knowledge there is no evidence that they do have a greater predictive success rate; in fact the impression is quite the reverse.

To improve data sources used in predictions would involve the public as individuals and groups in providing information on present and intended actions and aspirations which, for various reasons, they are unwilling to divulge with any degree of spontaneity. At least part of this unwillingness stems from a sense of isolation from the decision-taking which would (hopefully) be based on the information. If truly participatory decision-taking processes existed, after a due period of time to allow for cultural inertia, the inclination to provide adequate information from which to derive alternative policy plans might well emerge.

Better data require more democracy, not less.

DAVID R. COPE
Institute of Planning Studies, University of Nottingham, UK

Mistaken parameter

SIR,—Politicians and others have been criticised in the national press for the misuse of the word "parameters", but I have yet to see similar charges levelled against scientists who have much less excuse for this practice. During the past three years the tendency to describe components, factors, attributes and so on as parameters has become frequent in most scientific journals, covering agricultural research and related fields, published in the UK and abroad.

To anyone in the habit of using parameters with gay abandon I would recommend the definition given by Professor D. J. Finney in *An Introduction to Statistical Science in Agriculture*:

"Numerical values that serve . . . to identify a particular distribution curve, as a member of a whole family, are termed *parameters* . . ."

I would predict that if the term were restricted to numerical values, without any greater precision of meaning, the frequency with which it is used would be halved.

I. G. MONTGOMERIE

Mycology Section, Scottish Horticultural Research Institute, Dundee, UK

Electrifying demonstration

SIR,—The theoretical approach by Marble, MacVicar and Roberts (Correspondence, July 15, page 171) is not necessary. For many years, when teaching electrostatics to first year students, I used to stand on a 5 cm thick block of paraffin wax and charge myself up from a small Van de Graaf generator capable of giving 300,000 volts on open circuit. This made my hair stand on end in a convincing manner, and the sparks which I could get from a hand-held object to earth would certainly have ignited any explosive gas mixture.

Without wishing to compete with the above authors in numbers of significant figures, I must have reached between 100,000 to 200,000 volts with respect to Earth without harm—incidentally illustrating for the sake of the students the old GEC adage: "It's the amps, not the volts, wot kills yer".

J. H. FREMLIN

Department of Physics, University of Birmingham, UK

news and views

Eukaryotic mRNA : trouble at the 5'-end

from Beverly Griffin

UNTIL recently, the story of the 5'-termini of eukaryotic messenger RNAs seemed consistent and scientifically coherent. The evidence emerging seemed to confirm two separate conclusions. First, eukaryotic messenger RNAs contain 7-methylguanosine triphosphate, $m^7G(5')ppp(5')X$, at their 5'-ends. To use the terminology introduced by Rottman *et al.*, they are "capped".

This seemed to hold true for mRNAs from a large number of cells: mouse myeloma cells[2] and L cells[3,4], HeLa cells[5-7], Novikoff hepatoma cells[8], Erhlich ascites carcinoma cells[9], brine shrimp (*Artemisia salina*) cells[10] and silk fibroin (*Bombyx mori*) cells[11].

The same 5'-terminal ends were found for mRNAs from both DNA and RNA viruses, such as: reovirus[12], Newcastle disease virus[13], vaccinia virus[14,15], Sindbis virus[16,17]; avian sarcoma (ASV) virus[18,19], vesicular stomatitis virus[20,23], adenovirus[24], and cytoplasmic polyhedrosis virus[25]. They were also found in several plant viruses: tobacco mosaic (TMV) virus[26,27], alfalfa mosaic virus[28], and brome mosaic virus[29].

Translation requirement

Second, the presence of the 5'-terminal 7-methylguanosine (that is, the "cap") is required for translation. Two separate, but related, studies using mRNAs from reovirus or VSV showed that the removal of the cap resulted in a loss of translational activity in *in vitro* protein synthesising systems[30,31].

Following on these findings, other results on the how and when and where began to emerge. Aurin tricarboxylic acid (ATA) which inhibits translation, was found by Both *et al.*[31] also to inhibit messenger RNA methylation, leading them to suggest that the cap structure is important in messenger-ribosome complex formation, the absence of 7-methylguanosine markedly reducing formation of the complex. Rose[22] showed that messenger RNAs with 5'-triphosphate ends, but no 7-methylguanosine, were present in VSV but were not found on ribosomes;

he therefore postulated that capping might have a role in ribosome recognition of mRNA. Further corroboration and refinement came from the results of Roman *et al.*[32], who, studying mRNA binding in a fractionated wheat germ system, obtained results which suggested that the capped mRNA was essential for the interaction of the 40S ribosomal subunit with the 60S subunit to give the 80S ribosomal complex; in competitive inhibition studies, 7-methylguanosine-5' phosphate was found to have no effect on the formation of the 40S complex, but to inhibit the transition of the 40S ribosome-met-tRNA$_f^{Met}$ complex to the 80S ribosome complex. Similar experiments were carried out in other laboratories and similar results obtained[33-35]. In some interesting studies with model compounds, Both *et al.*[35] found that the presence of 7-methylguanosine triphosphate at the 5'-end of some ribopolymers was not in itself sufficient to ensure ribosomal binding. For example, capped polyriboadenylic acid did not bind to any significant extent, whereas capped polyribouridylic acid could form 40S complexes, and capped ribopolymers containing both A and U could form 80S complexes. Their results strongly implicated not only the 7-methylguanosine in ribosome binding, but also pointed to the need for sequences rich in A and U for stable binding. It does seem important to note here that some ribopolymers such as $(A,U)_n$ bound stably to ribosomes (even though they did not contain a 7-methylguanosine cap). The dependence on the presence of adenosines and uridines for stable complex formation between eukaryotic mRNA and ribosomes makes the sequence of brome mosaic viral RNA fragment protected by eukaryotic ribosomes all the more interesting; Dasgupta *et al.*[29] found the sequence which precedes the translation start signal in this message to be $m^7G(5')-ppp(5')GUAUUAAUA$ . . . , a sequence notably rich in A and U.

For a further probe into the role of capped mRNAs in protein synthesis, Shafritz *et al.*[36] studied the interaction

of mRNAs from VSV, EMC, and histone mRNA with purified initiation factors. They demonstrated complex formation between mRNAs and factors IF-MP, IF-M3, and IF-M2A (but none with IF-M2B). For inhibition studies, they used excess 7-methylguanosine triphosphate, and in only one case, that with IF-M3, was this nucleotide able to inhibit messenger-factor complex formation. Their conclusion from their studies was that initiation factor IF-M3 might be essential for the binding of certain mRNAs to ribosomes.

Exceptions

So far, so good. And now the trouble starts, and we come to the findings that begin to question the validity of the above observations as generalisations to all eukaryotic mRNAs. Two papers which appeared simultaneously say that polio viral mRNA is not capped by its 5'-terminal's being pUp[37,38]. Moreover, encephalomyocarditis (EMC) virus mRNA appears to contain no capped 5'-end[39], nor do the mRNAs from the plant virus, satellite tobacco necrosis (STNV) virus[33,40]. So not all eukaryotic mRNAs are capped.

Moreover, a paper has now appeared which says that translation does not require 7-methylguanosine; Rose and Lodish[41] were able to translate mRNAs of VSV in an *in vitro* system even after the terminal 7-methylguanosine had been removed. Thus, translation of some messengers in some protein synthesising systems may not require the terminal cap.

To add further to the uncertainties, Filipowicz *et al.*[42] find a cap-binding protein in extracts of brine shrimp embryos, but this protein which binds to mRNAs and is postulated to promote binding to ribosomes, is not IF-M3, nor indeed does it appear to be any of the other protein initiation factors.

It is clear from the paradoxes mentioned above, that different as well as better-defined experiments need to be carried out in order to reach definite conclusions about the structure and

function(s) of the 5'-end of eukaryotic mRNAs. With few exceptions, most of the studies to date have been so similar in kind that if a trap is there, it would have captured in turn almost every investigator. For example, most studies on mRNA have been carried out on species purified *via* the polyadenylated 3'-ends of the RNAs. But not all mRNAs contain poly(A), and who can yet say that poly(A) may not be added merely to "store" mRNAs until they are needed, and then removed to produce "fuctioning" messages? (Marbaix *et al.*,[43] point to the role of poly(A) 3'-terminal sequences as one of stabilising eukaryotic mRNAs.) Moreover, most studies on the function of the cap at the 5'-end of mRNAs have used mixtures of messenger RNAs and the wheat germ *in vitro* translating system. (There are a few exceptions to this criticism. For example, see Yang *et al.*[11] and Shafritz *et al.*[36] for use of the rabbit reticulocyte cell-free protein synthesising system.) Even after considering the exceptions, it is not clear whether one is justified in drawing firm conclusions about mRNAs from a single protein synthesising system. It seems from the data accumulated so far that studies on mRNAs may merely reflect the idiosyncracies of the individual *in vitro* translation systems. Moreover, with the mRNAs that fail to be translated, no real proof exists that the chemical procedures used to remove the capped ends have not damaged the RNAs in other ways. Rose and Lodish[41] say : "We initially had great difficulty recovering translatable VSV mRNA after periodate-oxidation and β-elimination, and clearly did not completely overcome the problem of degradation. Thus, we suggest that similar studies indicating a requirement for 7-methylguanosine in translation of cellular mRNAs should be more carefully controlled for possible non-specific effects of the chemical procedure".

If 7-methylguanosine survives as a requirement for translation for most mRNAs and in most protein synthesising systems (as well as in the cell?), it will be interesting to see whether mRNAs for which no cap has been found, such as those from poliovirus or EMC, have a 7-methylguanosine (as a modified base) anywhere near their 5'-termini.

More about the 5'-end

What else do we know about the 5'-ends of eukaryotic mRNAs? Most eukaryotic mRNAs, in addition to containing 7-methylguanosine triphosphate at their 5'-ends, also appear to contain 2'-O-methylated bases as the first, and frequently the second, nucleoside following the capped end. Some exceptions to this have been found, notably among the viral mRNAs. (For ex-

Salt on roadside verges

from Peter D. Moore

THE impact of disturbance upon the flora of roadside verges is a subject of considerable interest to conservationists and to environmental planners, for the habitat occupies a large area of our country and provides an opportunity for combining utility with aesthetic and scientific interests. In the main, interest has centred upon such effects as lead pollution from motor exhausts, but an additional stress imposed on these verges is provided by the salt which is applied to roads in winter.

The problem was recognised and discussed by Westing (*Phytopathology*, **59**, 1174; 1969) in America where, at the time of his survey, some 6 x 10[6] tons of salt were applied to the highways of the northern states during the winter. This figure was increasing by about a million tons every year. In the United States, applications of 6 kg m^{-1} along the side of the road are frequent. In Britain, Davison (*J. appl. Ecol.*, **8**, 555; 1971) estimated that applications of 3 to 4 kg m^{-2} were experienced by the roadside verges of Northumberland.

The presence of salt in the soil lowers the water potential and influences soil structure, which may make water less available and reduce aeration for the resident plants. The effects on trees have caused concern in the United States, where Westing (*Econ. Bot.*, **20**, 196; 1966) has described a decrease in the population of the sugar maple (*Acer saccharum*) caused by a combination of drought and salt applications. Low rainfall in

summer, of course, can result in reduced leaching and the persistence of salt applied during the winter.

In Britain, trees are not permitted to grow close to the edge of roads, so are not subjected to the greatest stress; this is felt instead by the characteristic grasses and herbs. Analysis of verge soils by Davison (*op. cit.*) showed that sodium levels in April could be ten times those in unaffected grassland, but usually fell to normal by September. Sodium concentrations in plant leaves were often five to ten times the normal level.

It may be expected that the continued application of salt to roadsides would result in the selection of more tolerant species and this has now been observed by Matthews and Davison (*Watsonia*, **11**, 146; 1976) in areas around Newcastle-upon-Tyne, Northumberland. In a survey of changes in roadside flora they describe reduced species diversity and the selection of resistant species. They also describe an invasion of maritime species, such as *Aster tripolium, Puccinellia maritima, Plantago maritima* and *Sueda maritima* up to 13 km from the sea. These halophytic species have been particularly effective in the colonisation of new (less than 3-yr old) roads. It is likely that these species have been brought in as seeds on the wheels of vehicles. It will be interesting to see whether the development of a saline roadside flora continues along our highways in a similar manner to the establishment of a railway flora in the last century.

ample, see the references to Newcastle disease virus, Sindbis virus and TMV virus mRNAs.) Structures of the type m[7]G(5')ppp(5')X^{m}pYp are now being called cap I structures and those of the type m[7]G(5')ppp(5')X^{m}Y^{m}pZp are being called cap II structures[44]. Perry and Kelley have carried out a very interesting study on the kinetics of methylation of the 5'-ends of mRNAs. Their results suggest that the 5'-terminal cap I structures of mRNAs are derived from capped heterogeneous nuclear RNA and are conserved during processing. Cap II structures, on the other hand, are not found in the cell nucleus, but arise by a secondary methylation after the mRNAs have entered the cytoplasm. This secondary methylation appears to be restricted to a particular subclass of mRNAs which have a high frequency of pyrimidine nucleosides at position Y (see above), and are among the more stable

mRNAs. Studying mRNA from cytoplasmic polyhedrosis virus, Shimotohno and Miura[45] have shown that methylation at the messenger 5'-end occurs first at the 7-position of guanosine, and only after chain elongation has proceeded for a few nucleotides is the 2'-position of the first base which follows the cap structure (adenine, in this case) methylated.

(The discovery in cell extracts of enzymes which cleave at the 5' terminal and remove the cap will no doubt help resolve some of the problems and conflicts now being encountered in studies on eukaryotic mRNAs. Three papers have appeared which report on the presence of such an enzyme activity in HeLa cells[7], in extracts from brine shrimp[46], and in cultured tobacco cells[47].

One of the early papers in this field, and probably the one which was most influential in bringing the capped ends

of mRNAs to the attention of the scientific public was the paper by Rottman, Shatkin and Perry[1] entitled "Sequences Containing Methylated Nucleotides at the 5' Termini of Messenger RNAs: Possible Implications for Processing". In one of the most important papers to appear and challenge the ideas widely quoted (and therefore beginning to be accepted as dogma) on the function of the cap in eukaryotic mRNAs, Rose and Lodish[41] say that their (conflicting) results "suggest that methylation is involved in synthesis or processing of (VSV) mRNAs into the correct poly(A)-containing mRNA species". And so, have we gone through so much only to come full circle?

I should like to conclude by making a few remarks. First, 7-methylguanosine is an extremely interesting molecule. One of the early papers on capped structures referred to the cap as a "bizarre" structure[2]. The better word might be "versatile". Haines *et al.*[48] found 7-methylguanosine to have a potentiometrically determined pK_a of 7.0. This means that at pH 7.0, it exists as an equilibrium between two ionic forms:
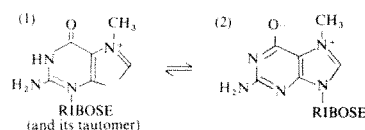
Lower the pH, and form (1) becomes more important. Raise the pH, and form (2) becomes more important. The implications of this are that the capped structure would be expected to be remarkably sensitive to localised pH changes within the cell.

In an unusually interesting paper in a too-neglected field, Seeman *et al.*[49] have studied sequence specific recognition of nucleic acids by proteins. For double-helical RNA (or DNA), their analyses lead them to conclude that whereas single hydrogen bonding interactions were inadequate to explain the specific protein-nucleic acid complexes they observed in model studies, pairs of hydrogen bonding interactions might play a role in protein-nucleic acid recognition. To return to the capped structure, within this structure (even ignoring any effects due to phospshate groups) would appear to be the requisite qualifications for specific interactions with proteins; cursory examination would suggest a possible stable interaction with, for example, the guanidinium group of an arginine molecule. With a slight change in pH, the tautomeric form of 7-methylguanosine could change and the complex be thereby strengthened, or weakened, depending on the nature of the change. The word versatile reflects the potential of the capped structure to respond in a very sensitive way to its environment; its environment in turn can confer on it the properties for self-regulation, whether in messenger binding or processing or transport.

The questions raised by the presence (or absence) of the cap at the 5'-ends of eukaryotic mRNAs have clearly stimulated considerable interest, and even that form of excitement which is so important to scientific research. But the tale is unfinished. And so, the reader is left with the opportunity to carry on with the story . . . .

1 Rottman, F., *et al.*, *Cell*, 3, 197 (1974).
2 Adams, J. M., and Cory, S., *Nature*, 255, 28 (1975).
3 Perry, R. P., and Kelley, D. E., *Cell*, 1, 37 (1974).
4 Perry, R. P., *et al.*, *Cell*, 4, 387 (1975); 6, 13 (1975).
5 Furuichi, Y., *et al.*, *Proc. natn. Acad. Sci U.S.A.*, 71, 1904 (1974).
6 Wei, C. M., *et al.*, *Cell*, 4, 379 (1975).
7 Nuss, D. L., *et al.*, *Cell*, 6, 21 (1975).
8 Desrosiers, R., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 71, 3971 (1974).
9 Bajszar, *et al.*, *Molec. Biol. Rept*, 2, 369 (1976).
10 Muthukrishnan, S., *et al.*, *J. biol. Chem.*, 250, 9336 (1975).
11 Yang, N-S., *et al.*, *Cell*, 7, 339 (1976).
12 Furuichi, Y., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 72, 362, 742 (1975).
13 Colonno, R. J., and Stone, H. O., *Nature*, 261, 611 (1976).
14 Wei, C. M., and Moss, B., *Proc. natn. Acad. Sci. U.S.A.*, 72, 318 (1975).
15 Urushibara, T., *et al.*, *FEBS Let.*, 49, 385 (1975).
16 Dublin, and Stollar, *Biochem. biophys. Res. Commun.*, 66, 1373 (1975).
17 Hefti, *et al.*, *J. Virol.*, 17, 149 (1975).
18 Furuichi, Y., *et al.*, *Nature*, 257, 618 (1975).
19 Stoltzfus, and Dimock, *J. Virol.*, 18, 586 (1976).
20 Abraham, G., *et al.*, *Nature*, 255, 37 (1975).
21 Abraham, G., *et al.*, *Cell*, 5, 51 (1975).
22 Rose, J. K., *J. biol. Chem.*, 250, 8098 (1975).
23 Moyer, S. A., *et al.*, *Cell*, 5, 59 (1975).
24 Sommer, *et al.*, *Nucleic Acids Res.*, 3, 749 (1976).
25 Furuichi, Y., and Miura, K., *Nature*, 253, 374 (1975).
26 Keith, J., and Fraenkel-Conrat, H., *FEBS Lett.*, 57, 37 (1975).
27 Zimmern, *Nucleic Acids Res.*, 2, 1189 (1975).
28 Pinck, *FEBS Lett.*, 59, 24 (1975).
29 Dasgupta, R., *et al.*, *Nature*, 256, 624 (1975).
30 Muthukrishnan, S., *et al.*, *Nature*, 255, 33 (1975).
31 Both, G. W., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 72, 1189 (1975).
32 Roman, P., *et al.*, *Nature*, 260, 359 (1976).
33 Hickey, E. D., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 73, 19 (1976).
34 Both, G. W., *et al.*, *J. molec. Biol.*, 104, 637 (1976).
35 Canaani, *et al.*, *FEBS Lett.*, 64, 326 (1976).
36 Shafritz, D. A., *et al.*, *Nature*, 261, 291 (1976).
37 Hewlett, M. J., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 73, 327 (1976).
38 Nomoto, A., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 73, 375 (1976).
39 Fellner, *et al.*, *3rd Int. Congress Virol. Abstr.*, 161 (1975).
40 Shih, D. S., and Kaesberg, P., *Proc. natn. Acad. Sci. U.S.A.*, 70, 1799 (1973).
41 Rose, J. K., and Lodish, H. F., *Nature*, 262, 32 (1976).
42 Filipowicz, W., *et al.*, *Biochemistry*, 73, 1559 (1976).
43 Marbaix, G., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 72, 3065 (1975).
44 Perry, R. P. and Kelley, D. E., *Cell*, 8, 433 (1976).
45 Shimotohno, and Miura, K., *FEBS Lett.*, 64, 204 (1976).
46 Filipowicz, W., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 73, 1559 (1976).
47 Shinshi, *et al.*, *FEBS Lett.*, 65, 254 (1976).
48 Haines, *et al.*, *J. Chem. Soc.*, 5281 (1962).
49 Seeman, N. C., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 73, 804 (1976).

Why do stars twinkle?

from B. J. Uscinski

ALTHOUGH the twinkling of stars is apparently a simple phenomenon, the theory which has been developed to explain it is anything but simple. The article by Jakeman, Pike and Pusey on page 215 of this issue of *Nature*, shows that it is now possible to test the validity of the new multiple-scatter theory which has evolved over the past 10 years. This is an important step forward, since the theory is essential not only for the study of optical scintillation, but also in investigating the scintillation of quasars and pulsars caused by the interplanetary and interstellar media respectively. Up till now it has not been possible to make satisfactory experimental tests of the new body of multiple-scatter theory since detection and recording techniques have not been sensitive enough.

The early 1960s saw the beginning of serious efforts to study the propagation of waves in media containing random irregularities of refractive index. One of the first attempts to deal with the case when the intensity fluctuations of the wave become large was the so called 'Method of Smooth Perturbations'. (Obukhov, *Izv. Akad. Nauk SSSR, ser. geofiz.*, No. 2, 155; 1953). This method became widely known, having appeared in a book by Tatarsky (*Wave Propagation in a Turbulent Medium* (trans. by R. A. Silverman), McGraw-Hill, New York, 1961), but it was soon realised that it was not valid for the case of multiple scatter (Barabanenkov *et al.*, *Usp. fiz. Nauk*, 102, 1; 1970). Subsequent methods based on repeated application of the Born approximation were able to deal adequately with the multiple-scattering of waves in media where the perturbation caused by a single irregularity is small (Shishov, *Izv. Vyssh. Ucheb. Zaved. Radiofizika*, 11, 866; 1968; Tatarsky, *Zh. Eksp. Teor. Fiz.*, 56, 2106; 1969; Beran and Ho, *J. opt. Soc. Am.*, 59, 1134; 1969). This came to be known as the Markov approximation, and such media as "weakly irregular" media.

The new theory is based upon a set of partial differential equations for the various moments of the field, and theoreticians generally agree that they are valid in the case of multiple scatter (Barabanenkov, *op. cit.*). The equations have several interesting features. First, the variables can be scaled in such a way that each equation contains only one parameter, which is the same for all the equations (Barabanenkov, *op. cit.*). Hence all weakly irregular media can be classified as regards their scattering properties in terms of a single parameter. Second, the level of intensity fluctuations, sometimes called the "scintillation index"

or the "normalised contrast", found by solving the equation for the fourth moment exhibits peaks or foci where values well in excess of unity occur in certain conditions (Dagkesamanskaya and Shishov, *Izv. Vyssh. Ucheb. Zaved. Radiofizika*, **13**, 16; 1970; Shishov, *Zh. Eksp. Teor. Fiz.*, **61**, 1399; 1971). Finally, a marked reduction in the scintillation index or normalised contrast is predicted when the signal is received over a wide frequency band.

It is indeed interesting that Jakeman *et al.* have been able to observe these last two effects. As the authors point out, earlier failures to observe the predicted effects have probably been due to instrumental insensitivity and averaging of the response. They do not necessarily indicate that the theory is wrong. The new detection techniques can now be put to an obvious and necessary use. Multiple scatter theory can be tested experimentally. This is essential since the theory is widely used as the basis for investigating stellar objects, as well as various aspects of both the interstellar and interplanetary media. It is also used when probing the atmosphere with laser beams.

The tests would have to be carried out in controlled conditions in a laboratory. As Jakeman *et al.* remark, some of the effects to be measured are heavily dependent on the model. If naturally-occurring irregular media are used it is difficult to measure the autocorrelation function and scale size of the irregularities. Moreover, there could well be a mixture of scale sizes present, and the statistics of the medium could be non-stationary. A stationary irregular medium with a single scale size and an autocorrelation function which can be determined could be synthesised in the laboratory from irregular plastic layers for example. In these conditions the new detection techniques would allow the predictions of multiple-scatter theory to be systematically verified, especially the focusing effect and the classification of media in terms of the single parameter.

In addition, the probability distribution of the wavefield could be determined for different scattering regimes. It has sometimes been asserted, with little theoretical justification and no qualification, that the logarithm of the amplitude is normally distributed. There seems to be more reason to suspect that the probability distribution is different for different regimes, and while being similar to a log-normal distribution in some circumstances is more like a Gaussian in others (Barabenkov, *op. cit.*).

Until controlled tests of this type are carried out, multiple-scattering theory will remain open to doubt and criticism. □

New limits on variability of fundamental physical quantities

from Paul Davies

ONE of the greatest mysteries of nature is why the various dimensionless ratios of fundamental physical quantities take the apparently arbitrary values that we find for them. Why, for example, should the proton mass M_p be 1836.1 times greater than the electron mass M_e and not some other, equally meaningless, number? Or why should the square of universal electronic charge, e^2, be 1/137.036 times the product of Planck's constant $\hbar$ and the speed of light c? If any old numbers will do, why not something elegant, like 1 or π?

Three schools of thought have arisen about these ratios. One supposes that the numbers are not arbitrary at all, but will emerge as a precise result from a future theory of matter, in the same way that the anomalous magnetic moment of the electron (0.00116 times the "normal" magnetic moment) may now be derived from quantum electrodynamics. The second does not attempt to explain these numbers, but points out that if they were very different from their observed values, it may not be possible for them to be observed—because the existence of living organisms (to do the observing) depends on many complicated processes more or less sensitive to these values.

The final school of thought is a more radical one, proposing that the numbers actually observed are indeed nothing special, because those numbers are not in fact constant, but are changing with time. Thus, the values 1836.1 or 137.036 for example, are just the 1976 values, and have no fundamental significance.

The reasoning which led to this idea arose in the 1930s from a study of cosmology. Several physicists and astronomers, most notably P. A. M. Dirac and A. S. Eddington, drew attention to the very large dimensionless ratio which one obtains by expressing the age of the universe in atomic units of some sort. This number is about 10^{40}. By apparent coincidence, 10^{40} is roughly the same ratio by which electric forces are stronger than gravitational forces between two atomic particles. Dirac argued that any fundamental theory of matter would be unlikely to throw up a number so preposterously large as 10^{40} so that perhaps the latter ratio is actually determined by the former, that is that there is a causal connection between these two large numbers. If that is so, then as the Universe grows older, and the first ratio increases, so must the force of gravity dwindle in relation to electromagnet-

ism. This could come about by a decrease of Newton's gravitational "constant" G, for example, or by an increase of the electronic charge e.

Fortunately, it is possible to provide observational checks on the hypothesis that some of these ratios may be time-dependent. This is because modern technology enables us to study matter in regions of the Universe which are so distant that the light from them has taken several billions of years to reach us. Therefore, we see those regions as they were when the Universe was only a fraction of its present age. Any variation in the numerical ratios which is based on a cosmological cause might therefore be expected to show up in the spectra from distant galaxies or quasars. The redshift of the spectral lines fixes the distance, and hence the age of the source according to Hubble's law of the expansion of the Universe. But if it is also possible to measure the fine structure of the spectral lines (a small splitting due to relativistic effects in atoms) then a value for the fine structure constant $\alpha = e^2/\hbar c$ may be determined. In 1967 Bahcall and Schmidt were able to fix, using this technique, the possible variation in α to be below about one part in 10^{12} per year, thus ruling out the theory that G is fixed and e^2 increases in proportion to the age of the Universe (about 1 part in 10^{10} per year).

Recently, not only fine structure, but hydrogen hyperfine structure has been detected in a very distant radio source. This hyperfine structure is an electromagnetic effect which depends on the small coupling between the proton and electron in the hydrogen atom, and its measurement yields a value for the ratio $\alpha^2 M_e/M_p$. In a recent issue of *Physical Review Letters*, three American astronomers, A. M. Wolfe, R. L. Brown and M. S. Roberts use the new observation of hydrogen hyperfine absorption to fix the variation of $\alpha^2 M_e/M_p$ to as little as two parts in 10^{14} over at least 35% of the age of the Universe (the "look-back" time of the source). They also confirm Bahcall and Schmidt's result by observing fine-structure in the magnesium spectrum from this source. These results taken together (and assuming the internal structure of the proton does not change) completely rule out any small power variation of α or M_e/M_p over cosmic time scales. Of course, Dirac's theory can only be ruled out by simultaneously providing a restriction on the variation of G, a much more

difficult task.

Although this is a negative result, the importance for theoretical physics of keeping the natural constants constant makes these occasional cosmological checks well worthwhile. In the absence of a theory of natural constants, the significance of their numerology will have to be left to the philosophers to ponder.

Herpes virus induction of F$_c$ receptors

from R. S. Kerbel

WHEN cells are infected with viruses or are transformed by oncogenic viruses, alterations in plasma membrane components often occur. The characteristics of these surface changes and how they may relate to the biological behaviour of the cells have been popular areas of study for both virologists and immunologists alike.

One example of how spectacular these changes can be concerns induction of surface receptors, specifically F$_c$ receptors. Over 12 years ago, Watkins noted that cultured HeLa cells were able to absorb antibody-coated erythrocytes and form rosettes after the cells were infected with herpes simplex type 1 virus (*Nature*, **202**, 1364; 1964). This was later confirmed by Yasuda and Milgrom who used various different cell lines derived from both man and monkey (*Int. Arch. Allergy*, **33**, 151; 1968).

We now know that a variety of cells, mostly lymphoreticular in nature, have surface receptors for the F$_c$ portion of IgG immunoglobulin molecules (F$_c$ receptors). These receptors can be detected by erythrocyte–antibody (EA) rosette formation. Thus, it seems that infection with herpes simplex virus (HSV) results in induction of F$_c$ receptors, a conclusion substantiated by Yasuda and Milgrom who showed that virus-infected cells could not absorb erythrocyte coated with F(ab')$_2$ antibody fragments, but only with the fully intact antibody molecules.

The gene which codes for the receptor seems to be virally coded, and not a repressed host gene switched on by viral infection (Westmoreland and Watkins, *J. gen. Virol.*, **24**, 167; 1974). Subsequent studies have shown that mammalian cells such as hamster embryo fibroblasts oncogenically transformed by inactivated HSV types 1 and 2 can also display F$_c$ receptors (Westmoreland, Watkins and Rapp, *J. gen. Virol.*, **25**, 167; 1974).

It was only natural to ask whether other members of the herpes family of viruses could also induce the appear-

ance of F$_c$ receptors in infected cells and an affirmative answer was recently provided by Furukawa *et al.* who found human fibroblasts displayed F$_c$ receptors after infection with human cytomegalovirus (*J. clin. Microbiol.*, **2**, 332; 1975), an observation also made by Rahman *et al.* (*J. Immun.*, **117**, 253; 1976), Keller *et al.* (*J. Immun.*, **116**, 772; 1976) and Westmoreland, Jeor and Rapp (*J. Immun.*, **116**, 1566; 1976).

Now that every type of cell thus far tested from over six different species has been shown to develop F$_c$ receptor sites after infection by at least two different herpes viruses, the possibility looms larger than ever that these receptors have a role in the natural history of herpes virus infections. It has been suggested by Westmoreland and Watkins and also by Costa and Rabson (*Lancet*, **i**, 77; 1975) that the development of F$_c$ receptor sites may play a part in conferring a biological advantage on the virus and therefore in maintaining the latent nature of herpes infections, since the coating of infected cells with IgG molecules or immune complexes by the F$_c$ region of the immunoglobulin molecules might protect the cells from the destructive effects of potentially cytotoxic antibodies and lymphocytes.

Lehner, Wilton and Shillitoe (*Lancet*, **ii**, 60; 1975) suggested an alternative way in which F$_c$ receptors may be involved in the latent and recurrent nature of HSV infections, namely "double" binding of anti-HSV IgG antibodies to virus-infected cells to induced F$_c$ receptors and HSV surface antigens, by way of their F$_c$ and F$_{ab}$ portions, respectively. The F$_c$ portion of the anti-HSV antibodies would then be unavailable for either complement binding or for F$_c$ receptors on effector cells capable of mediating antibody-dependent cell-mediated cytotoxicity. Lehner *et al.* also speculate that the spread of putative HSV-induced squamous cell carcinomas might also be facilitated by double-binding of F$_c$ receptors and HSV antigens by IgG antibodies on the surface of dividing carcinoma cells. If there is indeed a relationship between HSV and squamous cell carcinoma, it will be of interest to see whether F$_c$ receptors can somehow be exploited as a tumour cell marker in this particular case.

The *in vivo* significance of induced F$_c$ receptors on herpes virus-infected cells awaits further experimentation. For the moment, the emphasis is on the various practical *in vitro* considerations of the phenomenon. A particularly illuminating example is the recent work of Rager-Zisman, Grose, and Bloom (*Nature*, **260**, 369; 1976) who found that HSV-infected target cells could be non-specifically lysed by F$_c$ receptor-positive killer cells in the

absence of specific anti-HSV antibodies; the authors provided further evidence that this phenomenon might be mediated by the cross-linking of effector cells and target cells, both displaying F$_c$ receptors, by aggregated immunoglobulins or soluble immune complexes present in the serum of the culture medium. This supports the suggestion first made by Lehner *et al.* (*op. cit.*) that complexes of HSV and IgG antibodies might bind to F$_c$ receptors of both HSV-infected target cells and effector cells resulting in killing of the target cell.

Immune complex (or aggregated Ig)-dependent cell-mediated cytotoxicity may therefore represent one of the mechanisms by which F$_c$ receptor-positive lymphocytes from normal donors can spontaneously lyse tumour target cells *in vitro*, a phenomenon observed by several investigators (for example, Pross and Jondal, *Clin. exp. Immun.*, **21**, 226; 1975; Bean *et al.*, *Cancer Res.*, **35**, 2902; 1975) and which has often been a serious technical problem in specific cell-mediated microcytotoxicity assays. If so, the use of F$_c$ receptor-negative tumour target cells or culture media free of immune complexes and/or aggregated Ig should minimise or even prevent, in some cases, spontaneous lymphocyte-mediated cytotoxicity. □

Plant–microorganism interactions

from R. M. Cooper

A symposium on cell wall biochemistry related to specificity in host–plant pathogen interactions (organised by J. Raa and B. Solheim) was held at the most northerly university in the world in Tromsö, Norway on August 2–6, 1976. The proceedings of the meeting are to be published.

THE molecular basis of the specific parasitism or symbiotic associations of certain microorganisms with a plant species, or even variety, has fascinated but eluded scientists for many years. One such specific interaction is the so-called 'gene-for-gene' relationship between biotypes of plant pathogens and their hosts 'evolved' as a result of pathogens' overcoming new resistance genes introduced by man into crops. Similarly the symbiotic association between roots of legume species and strains of the soil bacterium *Rhizobium* (a major source of biologically-fixed nitrogen) can exhibit a marked degree of specificity. These two types of specificity comprised the

main theme of this excellent symposium.

Several examples of cell–cell recognition involve interaction of surface carbohydrates or glycoproteins of one cell and a protein of the other; such as mannoproteins in sexual mating in the yeast *Hansenula wingei* (C. Ballou, University of California), O-antigen in attachment of phage to bacteria, single terminal glycosyl residues as determinants of red blood cell types and glycoproteins in aggregation of sponge cells. Recently carbohydrates from cell wall surfaces of pathogenic fungi have been implicated in specificity by their ability to induce differentially antifungal compounds—phytoalexins—in resistant and susceptible host plant cells (Keen, *Science*, **187**, 74–75; 1975). P. Albersheim's group (including A. Ayers and B. Valent, University of Colorado) reported extensively on their findings with these 'elicitors'. Initially they isolated and partly characterised them from culture filtrates of *Colletotrichum lindemuthianum* (bean pathogen) and *Phytophthora megasperma* var. *sojae* (Pms) (soybean pathogen), and subsequently found them to be an integral part of the fungal cell wall as Pms elicitor could be released from walls using a technique designed to remove surface antigens from yeast. Purified Pms elicitor induced phytoalexin accumulation in soybean, at the same rate as during infection by Pms, when applied to plant tissue in extremely low amounts (10^{-13} mol). It is a small (less than 25 glycosyl residues) predominantly β-1,3-glucan, dependent for inducing activity on terminal sugar residues. Although elicitors may have a critical role in pathogenesis they do not seem to determine specificity, since : elicitors from different races of the two fungi appear identical; response of resistant and susceptible host varieties is similar; similar elicitors to Pms are found in two saphrophytic yeasts, which suggests that plants respond to microorganisms by recognising a few common elicitors. An exciting offshoot of this work could be the development of elicitors as nontoxic fungicides.

Agrobacterium tumefaciens induces tumours in a wide range of dicots and gymnosperms but not monocots. Prerequisites for infection involve wounding of the host and attachment of virulent bacteria within 15 min of inoculation; avirulent strains are unable to attach. J. A. and B. B. Lippincott (Northwestern University, Illinois) found that the bacterial binding site resided within the polysaccharide moiety of the lipopolysaccharide (LPS) from the outer cell envelope; this fraction when isolated from virulent but not from avirulent strains inhibited tumour formation by competing with bacteria for host attachment sites,

Counting globin genes

from Pamela Hamlyn

THE molecular biology of abnormal human haemoglobins, and the study of haemoglobin variations within populations, not only provides explanations for clinically observed symptoms of diseases, but also acts as a model for the understanding of the biochemical genetics of other inherited traits. Simplicity is not one of the virtues of this particular model system. There are six main globins, α to ζ, which combine in different ways to form the various haemoglobins. The ε and ζ chains are present only in the first few weeks of foetal life.

Nucleic acid hybridisation techniques have shown that there are two genes coding for α-globin, one for β and one for δ. And now, to complete the count of the major globin genes, the number of γ-globin genes has been directly estimated (Old *et al.*, *Cell*, **8**, 13; 1976).

Two types of γ-chains are found in normal foetal haemoglobin, differing only at amino-acid position 136 which can be either alanine or glycine. The simplest interpretation of this would be that there are two structural genes for γ-globin. However, the proportions of the two types of γ-globin change after birth, and also γ-chain mutants do not occur in the ratio predicted if there are only two fully expressed genes. These observations have led to the suggestion that there are two major and two minor γ-globin genes (Huisman *et al.*, *Ann. N. Y. Acad. Sci.*, **165**, 320; 1969).

Because there are several steps between the transcription of the gene and the appearance of the protein, direct measurements of the number of genes are necessary; deductions made from the protein ratios are not sufficient. This was achieved by Old and his colleagues using a probe for γ-globin

genes. Messenger RNA from foetal red blood cells was reverse-transcribed to provide cDNA complementary to α, β and γ mRNA. Hybridisation of these cDNAs to adult blood mRNA, which contains only α and β mRNA, left the γ cDNA unhybridised and it was isolated from the hybrids. The purified γ cDNA was then hybridised to total human DNA in cDNA excess and the amount of cDNA needed to saturate the cell DNA was measured. The hybridisation of β-globin cDNA to total cell DNA was used as a comparison since it is known that there is one β-globin gene. The results confirmed that the number of copies of the β-gene was one, and showed that there were two genes for γ-globin. The different ratios between the two γ-globins which occur during development must therefore be due to differential gene transcription, or mRNA translation, rather than to gene dosage.

The method used in this study to determine gene number relies on the measurement of the rate and final extent of hybridisation with the probe in excess. This has a greater accuracy than the previously used method (Ottolenghi *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **72**, 2294; 1975) which relied on the measurement of the $C_0t_{\frac{1}{2}}$ when cDNA was hybridised back to excess total DNA. This is because with the saturation method a doubling in the gene number results in a doubling of the final plateau value, whereas in the other method the estimate of gene number depends on a small difference in the $C_0t_{\frac{1}{2}}$ normally measured on a logarithmic scale. Using the saturation method of determining gene number it will be possible to determine the frequency of globin gene duplication in human populations.

showing that LPS determined the specificity of attachment. The host component of attachment was in the cell wall as purified walls also inhibited tumour formation by competing for bacterial sites. The active fraction was polygalacturonide which was inhibitory at less than 1 ng ml^{-1}. In contrast monocot cell walls gave little or no inhibition of tumour induction, thus the inability of the bacterium to infect this group may result from lack of attachment sites. This is an interesting contrast to results of T. Graham and L. Sequeira (University of Wisconsin) on infection of tobacco by *Pseudomonas solanacearum*, in which avirulent strains attach to mesophyll cell walls and become rapidly enveloped by material, whereas virulent strains remain unattached and multiply freely. Binding may be determined by host

lectins (glycoproteins with diverse properties, for example haemagglutination) which agglutinated avirulent but not virulent cells because of the possession of extracellular polysaccharide by the latter.

Many other aspects of host–parasite interaction were considered, including plant cell wall structure, induction, repression and modification of pathogens' polysaccharidases, lignification, and degradation of fungal cell walls by host enzymes; however, these were discussed in terms of general rather than specific resistance. Some perhaps unexpected results were presented by H. Mussell (Boyce Thompson Institute, New York) implying interactions between pectic enzymes (PG) and host cell walls in specificity for *Verticillium* wilt of cotton; purified fungal PG induced symptoms only in susceptible

cotton cultivars. This may be partly due to the ability of 'resistant' walls to bind rapidly more PG than 'susceptible' walls, and the release of extra proteins, including peroxidase and IAA oxidase (both of which may be involved in symptom expression) from susceptible walls by PG.

Intriguing results were reported on host–symbiont interactions in which recognition and attachment of compatible rhizobia is the first stage before infection and nodulation of root hairs. Most evidence implied that recognition may depend on interactions between root lectins and bacterial surfaces. For example, W. D. Bauer (C. F. Kettering Research Laboratory, Ohio) found that soybean lectin only binds to living cells of symbiotic strains of *R. japonicum* and not to non-symbiotic strains, although there were some exceptions reported by J. Paxton (University of Illinois) and F. B. Dazzo (University of Wisconsin). Albersheim's results implied that the active bacterial surface component is LPS, as lectins from four legume species only interacted with the LPS of their corresponding symbionts. These lectins also had enzymic activity (against LPS) which may be a prerequisite for infection as intact LPS of *E. coli* elicits phytoalexin production; thus lectins may provide not only a physical linkage between surfaces but could trigger other critical events including a localised swelling of bacteria and decrease in respiration rate. Dazzo extended these studies using immunochemical techniques and found a surface antigen unique to infective strains of *R. trifolii* which cross reacted with antigens on roots of its symbiont clover; mutation to loss of infectivity resulted in loss of the antigens. A surface protein from clover roots, trifoliin, is specifically able to bind to the polysaccharide antigen and to agglutinate infective cells, and this probably explains specific adsorption by providing a bridge between the clover and rhizobial antigens.   □

Antarctic ice and desiccation in the Mediterranean

from A. Hallam

ONE of the more spectacular results of the Deep Sea Drilling Project has been the confirmation that the deep Mediterranean basins are underlain by a thick layer of apparently shallow water late Miocene (or 'Messinian') evaporite deposits of the type previously known from neighbouring land areas such as Italy. These evaporites are covered by Pliocene muds containing microfossils indicative of deep water conditions. For several years widely differing views have been put forward to account for such a peculiar deep sea occurrence. Hsu and his associates have argued that the northward convergence of Africa on Europe effectively cut off, about 6.5 million years ago, a series of deep topographic depressions in which, following desiccation, halite and calcium sulphate salts were precipitated in shallow salt pans many hundreds of metres below ocean level. Following the opening of the Straits of Gibraltar in the early Pliocene, Atlantic waters cascaded in to fill the basins and restore more or less normal oceanic conditions. Another school of thought rejects this interpretation, maintaining that the deep sea evaporites were originally deposited close to ocean level and occupy their present position as a consequence of substantial post-Miocene tectonic subsidence.

One means of deciding between these alternative hypotheses is to examine microfossils from deep sea core samples taken from the normal marine deposits directly below the evaporites. If these fossils are closely related to groups known to live at great depths in the present world ocean then it becomes difficult to argue that a shallow Miocene sea dried up and the present Balearic, Tyrrhenian, Ionian and Levantine basins were created by subsequent tectonic subsidence. In the proceedings of a micropalaeontologists' symposium held in Tunis (special issue of *Palaeogeography, Palaeoclimatology, Palaeoecology*, **20**, 1/2; July 1976), Benson reports on ostracods and foraminifera obtained from sub-evaporite core material in the Balearic Basin, which appear to indicate unequivocally water depths in excess of 2,000 m. Similar microfossils have been found in contemporary sediments in the Guadalquivir Basin of southern Spain and suggest a continuous oceanic passage through that region from the Atlantic to the western Mediterranean in pre-Messinian times.

Most of the symposium was devoted to discussing the profound effect that the so-called Messinian 'salinity' or 'desiccation crisis' in the Mediterranean had on marine faunas. The pre-Messinian microfaunas, such as ostracods and planktonic and benthonic foraminifera, were very similar if not identical to Atlantic types. Mass extinction took place during the desiccation catastrophe but repopulation from the Atlantic source commenced in the early Pliocene once normal marine conditions were restored. The effects of the catastrophe on the evolution of individual species could well be intriguing but has still not been fully evaluated.

Most scientists who have considered these extraordinary events in the geologically recent history of the Mediterranean have not sought for a world context, but there appears to be good support for Berggren's suggestion that the Messinian evaporitic phase corresponds with a global fall of sea level. Ecological analysis of foraminifera suggests a 40 m drop, which is in remarkably close agreement with the eustatic fall inferred from oxygen isotope data, due to a phase of growth of the Antarctic ice sheet at this time. It seems as though the structural and topographical situation of the Mediterranean region was such that a comparatively modest drop in sea level was sufficient to isolate it substantially from the world ocean, and well over a million years passed before a further eustatic rise allowed a normal connection to be renewed.

A hundred years ago

THE Association finds a fitting home in Glasgow, which has few rivals either in earlier or later scientific reputation. The force of long-continued scientific traditions, added to the present encouragement given to science, and I must also say, to the nearness of the finest holiday localities, makes this one of the most brilliant of recent meetings. Not only is the total number of members and associates attending very high, over 2,700, but the true chiefs of science are present in great strength. It cannot be said that the Association itself is this year at all below its high aims. The majority of papers are really scientific, and do not emasculate the truth in the effort to popularise it. Discussions have been very interesting, judging from the perseverance with which they have been listened to. The reception given by the people of Glasgow is worthy of the city, although it is possible that in the details and refinements of arrangement, Bristol excelled. This was especially manifested in regard to some of the excursions. But it is evident that the very best efforts of the north have been put forth in every way, and the general result is undeniably successful. The charming situation of the University Buildings, in which all the sections but one hold their meetings, is a very great advantage.

From *Nature*, **14**, September 14, 425; 1876.

articles

Calculations for cancer radiotherapy with pion beams

J. E. Turner, R. N. Hamm & H. A. Wright*

Clinical trials with negative-pion beams for cancer radiotherapy will begin soon at four special facilities in the USA, Canada, and Switzerland. Negative pions are elementary particles which can be produced in high energy accelerators and which have properties that make them seem promising as a new way of treating tumours. The complex interactions that pions undergo in matter make it difficult to calculate dose patterns produced in a body irradiated by a pion beam and to solve that problem a unique combination of research in high energy physics and medicine is needed. Monte Carlo computer codes developed for radiation shielding and radiation protection studies have been adopted for use in dosimetry and treatment planning with negative pions. In this paper we briefly compare the properties of negative pions with those of other types of radiation and describe some calculations that we have made to help assess the potential of negative pions for cancer radiotherapy.

THE question of negative pions for treating cancer was first raised by Fowler and Perkins in 1961[1]. Although negative pions have been available in some physics laboratories since the late 1940s[2], they have not been produced in sufficient intensity for radiotherapy and the large doses needed for treatment of tumours would have required impractically long irradiation times. Today, however, the first studies of pion radiation therapy are beginning in new facilities at the Los Alamos Scientific Laboratory, at Stanford University, at the Swiss Institute for Nuclear Research, and at the Tri-University Meson Facility in Vancouver[3], as adjuncts to primary research programs in physics. Dose rates at these installations are many times greater than those available before.

Because of the expense and difficulty of producing pions in the laboratory, the number and type of experiments which can be performed to assess their potential for cancer therapy are limited. Furthermore, the complex interactions that pions undergo do not lend themselves to analytical calculations. But the availability of high speed computers with large core memories makes it possible to simulate, on the computer, the traversal of a beam of pions in tissue, and programs have been developed specifically to help in studies with negative pion beams in cancer radiotherapy.

Properties of the negative pion related to therapy

The existence of the pion was predicted on theoretical grounds by Yukawa in 1935 to account for the strong, short range nuclear forces that hold atomic nuclei together[4]. Pions, which have a rest mass 280 times that of the electron, can be charged or neutral and are rapidly destroyed through strong interactions with matter and through radioactive decay. They are produced in collisions of high energy particles. Although there were no machines in the 1930s capable of accelerating

particles to the required energies of several hundred million eV, it was recognised that sufficiently energetic cosmic rays bombard the Earth. The presence of the pion in cosmic radiation was confirmed in 1947 (ref. 5).

There are almost 100 known elementary particles and anti-particles, all of which might, in principle, be used for radiotherapy. Negative pions are now attracting attention because of their special physical properties. Like other charged particles, the depth of penetration of negative pions in a target can be varied by varying their energy, and they deliver a relatively high dose when they stop at the end of their track—in the so-called Bragg-peak region. Unlike most other charged particles, however, a stopped negative pion is attracted by a positive atomic nucleus and interacts strongly with it through the nuclear force. The nucleus absorbs the pion and literally explodes with 140 MeV of energy converted from the rest mass of the pion. Figure 1 shows an early cloud-chamber photograph of charged-particle tracks around a nitrogen nucleus that has captured a stopped negative pion. Nuclear capture at the end of its range is the special property that makes the negative pion of particular interest for therapy.

The fragments produced by the capture give additional dose to the matter surrounding the capture site (Fig. 1). The biological effectiveness of this dose is enhanced by the fact that a substantial fraction of it is deposited at high linear energy transfer, since high LET radiations are generally more effective in producing biological effects. Calculations indicate, for example, that some 15–20 MeV of energy is deposited in tissue within 0.2 cm of the capture site by particles with LET greater than that of a 2-MeV proton (170 MeV cm^{-1})[6].

Furthermore, compared with low-LET, the effectiveness of high-LET radiation in killing cells is not reduced as much by the absence of oxygen. A rapidly growing tumour can presumably outgrow its blood supply and become hypoxic. In these conditions, it is well known that the dose of X rays or other low-LET radiations has to be two or three times higher to achieve the same degree of biological damage as in the aerobic state. Pions, in contrast, seem to offer a much more favourable

*Health Physics Division, Oak Ridge National Laboratory, Oak Ridge, Tennesse 37830.

Fig. 1 Cloud-chamber photograph of negative pion capture by nitrogen. The pion enters the picture from below and follows a curved path in the magnetic field through which it passes. The star at the end of the track is produced by two alpha particles and two protons. Four neutrons are also produced in the star but leave no visible tracks[7].

"oxygen enhancement ratio". Hypoxic cells can be killed almost as effectively as oxygenated cells with high-LET radiation (for more detailed discussion see ref. 8 and references cited there).

Basis for clinical trials

In using radiation to treat a tumour, the objective is to deliver a large dose over a specified volume in the body and as little as possible elsewhere. In reality, however, the dose patterns obtainable are limited by the physics of the interaction of radiation with matter.

Fig. 2 shows a comparison of dose patterns in water along the axes of beams of ^{60}Co γ rays, 22-MeV X rays, protons and negative pions (based on Raju and Richman[8]). The levels of each kind of radiation are adjusted to give the same average dose over the interval between 7.5 and 12.5 cm. The energy spectra of the proton and pion beams were additionally adjusted to give a constant dose over this interval. The dose from ^{60}Co γ rays builds up in the first few millimetres in the absorber as the number of secondary electrons increases. The dose decreases beyond about 0.5 cm as the radiation is further attenuated and scattered. The curve for the more energetic X rays has a maximum near 3 cm and falls off less sharply with increasing depth. In contrast to the neutral radiations, energies of the charged protons and pions can be controlled so that they stop in the given region, where they deposit the highest dose. Whereas a stopped proton gives no further dose, a stopped negative pion produces additional dose through nuclear capture. The advantage gained by negative pions over protons is illustrated by the fact that the pion dose is considerably smaller between the surface of the target and the region between 7.5 and 12.5 cm.

The factors described so far constitute the theoretical basis for the research and investment in pions in radiotherapy. Medical rationale and plans for clinically treating specific types of neoplasms with negative-pion beams have also been developed[9-11].

Computer codes for use in pion dosimetry

Several computer codes[12-15] are available for making calculations for pion beams incident on a patient or tissue-equivalent phantom. Calculations of pion-dose distributions pose several problems not encountered for other charged particles or conventional X or gamma rays. Because of the violent reactions that pions undergo, dose in the phantom does not readily lend itself to a description by means of simple analytical formula. Knowledge of the dose pattern from one beam-target configuration is of limited usefulness in predicting dose under different conditions. The variety of secondary particles produced— each with its own special properties—also complicates any analytical representation.

We have used a Monte Carlo technique to simulate on a computer the statistical distribution of events which occur to the individual particles in the phantom. By treating several thousand incident particles and the secondary particles they produce, one can thus obtain a detailed physical picture of what occurs in actual experimental conditions. The reliability of this picture depends directly on the reliability of the experimental data used in the calculations.

PION-1 was designed to treat all processes which pions undergo with sufficient accuracy to simulate real pion transport for radiotherapy applications. (It was developed in the Health Physics Division at ORNL and is at present restricted to pions, muons and electrons, with energies up to 125 MeV in materials consisting of H, C, N, and O. Other computer codes, developed in the Neutron Physics Division, are designed to transport pions, muons, protons and neutrons with energies up to hundreds of GeV. These codes use an intra-nucleus cascade model to calculate nuclear reaction products. PION-1, which makes use of some data calculated by the Neutron Physics codes, runs considerably faster than the more general, multipurpose codes.) Experimental values of the cross sections have been used to determine pion–nucleon interactions. These values can be revised or supplemented should data become available. All relevant pion nuclear reactions are programmed into the code: pion absorption, one- and two-nucleon knockout, and charge exchange. When an incident pion enters the phantom, the Monte Carlo technique is used to determine its fate, on the basis of these cross sections. As the pion penetrates the phantom, it undergoes multiple Coulomb scattering and also slows down in accordance with the stopping-power formula for charged particles. Range straggling is also included. When a nuclear reaction occurs, the code selects the type of reaction and the

Fig. 2 Examples of dose patterns in water along the axis of beams of ^{60}Co γ rays (1.33 and 1.17 MeV), 22-MeV X rays, protons (P), and negative pions (π^-).
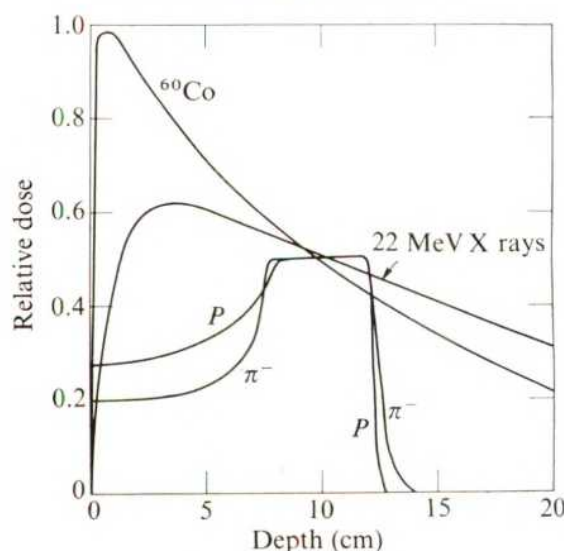
energies and initial directions of travel of any secondary particles produced. The secondary particles are also followed individually in the calculations. The use of numerical data, rather than nuclear models, allows the program to process particles rapidly.

For analysis, the phantom is divided into a number of sub-volumes. A record is kept in the program of each particle that traverses or stops or starts in every subvolume. The energy deposited, the LET of the radiation, and other physical quantities of interest are tabulated—the Monte Carlo method gives complete physical detail. Parameters for use in cell survival models can also be tabulated. Such data would be prohibitively expensive and difficult to obtain experimentally.

Within the context of negative-pion therapy, PION-1 is versatile. No special assumptions about the beam and phantom geometries are necessary. Pions can have any pattern of incidence on the phantom and any energy distribution up to 125

Fig. 4 Isodose contours for pion beam incident on soft-tissue target.

Fig. 3 Total dose, calculated (smooth curve) and experimental (circles), for CERN negative-pion bema. Radius of the detector was 3 cm. Other curves show various calculated contributions to the total dose. Labels on the curves indicate doses from the following sources: pion ionisation (π^-), heavy particles and protons generated from interactions of pions in flight (h+p), muon contamination (μ_c^-), electron contamination (e^-), heavy particles and protons generated from interactions of stopped pions (stars), and neutrons (n). The beam was assumed to be composed of $63\%\,\pi^-$, $23\%\,e^-$, and $14\%\,\mu^-$ with mean momentum 176 MeV/c(energy 85 MeV) and momentum spread 2.0%.

MeV. (The program can be extended to higher energies, if desired.) Muons and electrons can be present. The phantom can be composed of water, tissue, or other tissue-like material and can have regions of bone, lungs, air cavities, and so on.

The code runs rapidly. In a typical calculation, 10^4 incident pions and all the secondary particles they produce are processed completely. The statistical data are then excellent for plotting depth–dose and isodose contours and for compiling LET distributions and parameters for cell survival. The complete run takes about 10 min on the IBM 360/91. The processing of neutrons requires by far the most time. Without neutrons, the running time for 10^4 pions and the other secondary particles is less than three minutes. In many applications we include neutrons in only a fraction of cases.

Comparison of measurements and calculations with PION-1

Ideally, any calculation technique should be substantiated through detailed experimental checks of its component parts. Because of the complexity of pion interactions, detailed veri-

fications are not feasible. Many of the experiments would be costly and time consuming, while others would be prohibitively difficult.

Some direct comparisons have been made between measured doses and doses calculated with PION-1. Figure 3 shows one such comparison made for an experimental negative-pion beam from the 600-MeV synchrocyclotron at the European Centre for Nuclear Research (CERN) in Geneva, Switzerland. This work was carried out as a collaborative effort between the Health Physics Division at the Oak Ridge National Laboratory and the Health Physics Group at CERN[12]. The top curve, marked 'Calc', represents the dose calculated as a function of depth in a water target used in the experiment. The open circles, designated 'Exp', represent the values of the dose measured experimentally. The computations were performed for a beam mixture of 63% negative pions, 23% electrons, and 14% negative muons. A Gaussian momentum distribution was assumed with a mean value of 176 MeV/c(energy of 85 MeV) and a spread of 2% (that is, half of the particles have momentum within ±2% of the mean). The beam, which was non-uniform, was approximately elliptical in shape with major and minor axes of length 7 and 6 cm. The circular cylindrical detector of radius 3 cm, placed at different depths along the beam axis in the experiment, was simulated in the computations. The calculated individual contributions, which add up to the total dose 'Calc', are also shown in Fig. 3. The curve marked 'π^-' gives the dose that results from the direct ionisation of atoms by the pions in slowing down. The curve 'h+p' shows the dose from heavy recoil particles (atomic mass number >1) and recoil protons produced by pions in flight. (About 15% of the incident pions

Fig. 5 Isodose contours for pion beam incident on soft-tissue target with bone cylinder between 1 cm and 3 cm in the cross-hatched region.
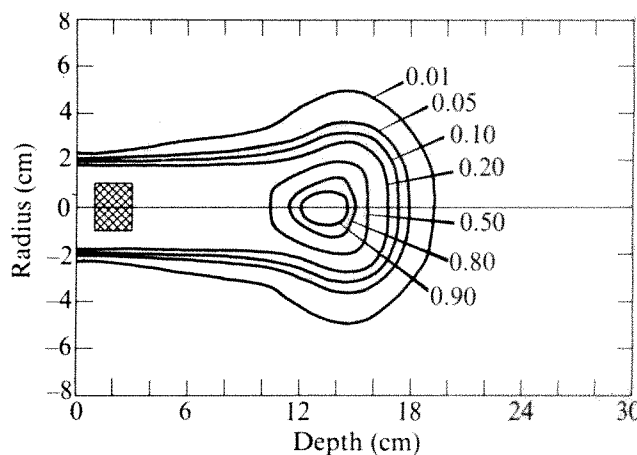
react with nuclei in the target before they reach the end of their range.) The local dose in the stopping region from the charged particles produced by pion capture is shown by the curve labelled 'Stars'. Neutrons, which occur primarily as a result of pion capture, give the dose shown by the curve 'n'. The doses from the muons and electrons in the experimental beam are shown by 'μ_c^-' and 'e$^-$'.

This comparison indicates that the dose from a beam of negative pions can be calculated from the known physics of pion interactions with some degree of confidence. Uncertainty in the calculated dose seems to be comparable to the reproducibility of the experimental measurements. Although direct experimental comparisons with most of the individual contributions shown in Fig. 3 have not been made, the curves are consistent with known physical data. The calculated number and energy spectra[16] of the particles produced by stopped-pion capture by oxygen have been compared with measurements[17] and seem to be in reasonably good agreement.

Dose patterns with negative pion beams

Of more direct interest for therapy are the dose patterns expected in tissue irradiated by different pion beams. Figure 4 shows calculated isodose contours in a tissue target of unit density from a parallel, uniform circular beam (radius 2 cm) of negative pions, having a Gaussian momentum distribution with a 2% spread and a mean energy of 67 MeV. The pions have a mean range of 15.0 cm. The beam is incident from the left, the axis being the line marked 0. The contours have cylindrical symmetry about the beam axis and apply to any plane containing that axis. The levels are shown as fractions, varying from 90% to 1% of the peak dose, 1.79×10^{-8} rad per incident pion, which occurs at a depth of 14.5 cm. The dose is not confined completely to the region directly intercepted by the beam. The beam spreads as it traverses the target because of multiple Coulomb scattering and nuclear reactions, which scatter particles outside the immediate beam area. The neutrons produced can deposit dose many centimetres away from the beam axis. Repeating the calculations but ignoring the neutrons gave virtually the same contours for the levels of 5% of the peak dose and greater. The peak dose itself was reduced, however, to 1.71×10^{-8} rad per pion. The lateral spread of the 1% contour from the beam axis at a depth of 15 cm changed from 4.9 to

4.4 cm when neutrons were neglected.

The presence of air cavities, bone and other regions of different densities in the body can greatly affect dose patterns. Figure 5 shows an example of dose contours when bone (density, 2 g cm^{-3}) is present. The bone is in the form of a right circular cylinder of radius 1 cm, centred on the beam axis between the depths 1 cm and 3 cm, as indicated by the cross hatching in Fig. 5. The same pion beam was used in the calculations of Figs 4 and 5. The peak dose with the bone present is decreased to 1.46×10^{-8} rad per incident pion and occurs at a depth of 12.6 cm as compared with 14.5 cm before. The bone increases the amount of multiple Coulomb scattering and some additional nuclear reactions, causing the dose pattern to be more diffuse in Fig. 5. A systematic study of the effects of tissue inhomogeneities has been completed recently[18,19].

Another complication, which must be assessed in using pions for radiotherapy, results from the neutrons produced from pion capture. The neutrons will travel away from the capture site in a tumour and irradiate healthy tissues elsewhere in the body. The neutron dose delivered to a sensitive organ in the vicinity of a tumour may limit the therapeutic dose that can be given to the tumour. We are now calculating the dose contributions from neutrons for a variety of pionbeam configurations and tumour sizes.

1 Fowler, P. H., and Perkins, D. H., *Nature*, **189**, 524–528 (1961).
2 Gardner, E., and Lattes, C. M., *Science*, **107**, 270–271 (1948).
3 Rothenberg, M. S., *Phys. Today*, **28**, No. 2, 17–20 (1975).
4 Yukawa, H., *Proc. phys.-math. Soc., Japan*, **17**, 48–57 (1935).
5 Lattes, C. M. G., Occhialini, G. P. S., and Powell, C. F., *Nature*, **160**, 453–456 and 486–492 (1947).
6 Dutrannois, J. R., Hamm, R. N., Turner, J. E., and Wright, H. A., *Phys. Med. Biol.*, **17**, 765–770 (1972).
7 Ammiraju, P., and Lederman, L. M., *Nuovo Cimento*, **4**, 283–306 (1956).
8 Raju, M. R., and Richman, C., *Curr. Top. Rad. Res. Quart.*, **8**, 159–233 (1972).
9 *Proceedings of the Conference on Particle Accelerators in Radiation Therapy, 1972* (National Technical Information Service, US Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22151).
10 Kligerman, M., *Eur. J. Cancer*, **7**, 191–194 (1971).
11 Kaplan, H. S., *Eur. J. Cancer*, **7**, 195–198 (1971).
12 Turner, J. E., *et al.*, *Rad. Res.*, **52**, 229–246 (1972).
13 Chandler, K. C., and Armstrong, T. W., *Nucl. Sci. Eng.*, **49**, 110–111 (1972).
14 Li, G. C., Boyd, D., and Schwettman, H. A., *Phys. Med. Biol.*, **19**, 436–447 (1974).
15 Wright, H. A., Hamm, R. N., and Turner, J. E., *ORNL-TM Report* (in the press).
16 Guthrie, M. P., Alsmiller, R. G., Jr, and Bertini, H. W., *Nucl. Inst. Meth.*, **66**, 29–36 (1968).
17 Fowler, P. H., and Mayes, V. M., *Proc. phys. Soc.*, **92**, 377–389 (1967).
18 Hamm, R. N., Wright, H. A., and Turner, J. E., *ORNL-TM-5088 Report* (Oak Ridge National Laboratory, October 1975).
19 Hamm, R. N., Wright H. A., and Turner, J. E., *Phys. med. Biol.* (in the press).

The origin of deuterium

Richard I. Epstein

Center for Astrophysics, Harvard College Observatory and Smithsonian Astrophysical Observatory

James M. Lattimer

Department of Astronomy, University of Texas at Austin and Enrico Fermi Institute, University of Chicago

David N. Schramm

Enrico Fermi Institute, University of Chicago

General nuclear constraints are used to show that deuterium is most likely of pregalactic origin. Big-bang nucleosynthesis is the most plausible source for significant amounts of this isotope, but other, more speculative, sources are not ruled out.

THE amount of deuterium which is formed during the initial phase of a standard hot big-bang universe is a sensitive function of the mean baryon density ($X_D \propto \rho^{-5}$ over the range of interest; X_D is the deuterium mass fraction and ρ is the mean baryon density; (see ref. 1). If big-bang nucleosynthesis were the only significant source of deuterium, the amount of deuterium

produced by this event would have to be at least as great as that which is currently observed, $X_D \sim 2 \times 10^{-5}$ (refs 2 and 3). The mean mass density of the present Universe would then be $< 6 \times 10^{-31}$ g cm^{-3} and hence less than the Einstein–de Sitter critical density $\rho_c = 3H^2/8\pi G = 4.5 \times 10^{-30}(H/50)^2$ g cm^{-3} (H is the Hubble expansion parameter in km s^{-1} Mpc^{-1}). In the framework of Friedmann cosmologies with cosmological constant $\Lambda = 0$ this would mean that the Universe is open and forever expanding[4,5]. Conversely, if the mean density of matter in the Universe is appreciably $> 6 \times 10^{-31}$ g cm^{-3}, the current deuterium abundance could not have been produced by nucleosynthesis during the early phase of a standard universe. These conclusions are insensitive to quite substantial deviations

from the standard cosmological models[6], presuming also that the bulk of the observed [4]He was produced in the big bang.

These statements provide good reason for investigating other possible sources of deuterium. If an exhaustive study yields no physically self-consistent or at least plausible models for deuterium production, then this failure lends support to the notion that deuterium is of cosmological origin. On the other hand, if an eminently plausible source of deuterium production or enhancement could be found, then the current deuterium abundance would not necessarily be related to the physical conditions in the early Universe and hence could not be considered to be an important cosmological discriminator.

We examine here various non-cosmological mechanisms which might be capable of producing significant amounts of deuterium. First, the general nuclear constraints which must be satisfied by all viable models for deuterium production are summarised. Specific models for deuterium sources are then discussed and critically analysed. The net result of this investigation is that very severe restrictions can be placed on mechanisms for producing deuterium. In fact, it seems unlikely that objects or events which are currently known or inferred to have existed in our Galaxy or in well-observed extragalactic objects could be capable of producing the observed deuterium abundance.

Overview

Deuterium formation can occur either through synthesis reactions in which free nucleons fuse together or by means of spallation or photodisintegration reactions in which energetic particles disrupt more massive nuclei, thereby ejecting deuterons. All models of deuterium production invoke one or both of these processes.

Synthesis models run into difficulty because the conditions which are needed for the making of deuterium are normally the same as those for its rapid destruction. Deuterium can be synthesised by [1]H(n,γ)D reactions if free neutrons and protons are able to interact. Free neutrons are produced at high temperatures, when heavier nuclei thermally dissociate, but any deuterium which is so formed is liable to be consumed in thermonuclear reactions. Deuterium could survive a hot explosive event only if the hot matter is dispersed and/or cooled rapidly compared with the time which is required for deuterium destruction. It is shown below that an object of even moderate mass ($\gtrsim 10^{-7} M_\odot$) cannot expand rapidly enough for significant amounts of deuterium to survive. Free nucleons may also be produced by the passage of a strong shock wave[8-10]. In this type of process the gas cools quickly enough for the deuterium to escape destruction, but models which invoke shock waves for deuterium production are faced with a number of other difficulties[11-15] which we also discuss below.

The possibility that a disrupting neutron star could provide a cool flux of particles has been examined (refs 16, 17 and J. M. Latimer and D. N. Schramm, unpublished), but it was found that this type of event cannot simultaneously provide both free neutrons and free protons, and hence cannot produce deuterium. Cameron and Truran (unpublished) have suggested that if hot neutron-rich material could be ejected by MHD processes from a collapsing stellar core[30,31], deuterium production could result. The constraints on this type of model are discussed below.

Spallation

Models which involve the production of deuterium by spallation reactions are faced with the problem that under most conditions the concomitant production of other light nuclei or γ rays is much too rapid. This is an especially strong constraint because deuterium is destroyed by thermonuclear reactions more easily than any other stable nucleus; any subsequent processing of the source material can only decrease the abundance of deuterium relative to other nuclei.

The interaction of energetic particles with normal population-I composition forms deuterium mainly by p + α reactions.

Other light nuclei, such as Li, Be and B, are formed by the interaction of protons or α particles with C, N, or O nuclei and by α + α reactions, and γ rays are formed by high energy p + p reactions producing π° which subsequently decay to two γ rays, each with an energy ~ 70 MeV. By considering the nuclear reactions which can occur in a hydrogen–helium gas mixture only, we can obtain very strong constraints on deuterium source models; these constraints are applicable to deuterium production in the present epoch and in pregalactic but post-helium synthesis epochs. (Deuterium production from spallation is impossible before helium synthesis.)

The rate at which a particle of type i is produced by the interaction of particles j and k is

$$\frac{dY_i}{dt} \propto Y_j Y_k \, \sigma_{i,jk}(E)$$

Here Y represents the number fraction of a particle (Y_i is the number of particles of type i per nucleon; $Y_i = X_i/A_i$ where A_i is the atomic weight), $\sigma_{i,jk}$ is the production cross section and E is the energy per nucleon of the incident particle.

The relative production rate, which is defined by

$$R_i \equiv \frac{1}{Y_{i,\text{present}}} \frac{dY_i}{dt} \qquad (1)$$

gives the rate at which the abundance of a particle approaches the present value. When two nuclei have the same value of R, their abundance ratio at the source is the same as their present cosmic abundance ratio.

The relative production rates for deuterium, [6]Li, [7]Li and γ rays are shown in Fig. 1 as a function of the energy per nucleon of the incident particle. The system is assumed to contain initially only hydrogen and helium with Y_α/Y_P taken to be 0.07 in the energetic particle flux and in the ambient medium. The present densities of the other particles relative to deuterium are $Y_{6\text{Li}}/Y_D \simeq 10^{-5}$, $Y_{7\text{Li}}/Y_D \simeq 10^{-4}$ and $Y_\lambda/Y_D \simeq 10^{-3}\rho_c/\rho$ (refs 18, 19). The last ratio is determined by using the number of photons in the diffuse background with energies > 70 MeV and assuming that the galactic deuterium abundance $Y_D = 10^{-5}$ is universal. It is important to note that in the present epoch there is far more deuterium than there is [6]Li, [7]Li or γ rays. Any class of events which produces even a relatively small amount of these particles cannot be the dominant source of the galactic deuterium.

Fig. 1 The rates at which abundances approach their present values as a function of the energy per nucleon of the incident particle. See equation (1) and text. Data are from refs 19, 28, 29 and C. King, private communication. Dashed lines are extrapolations from known data.
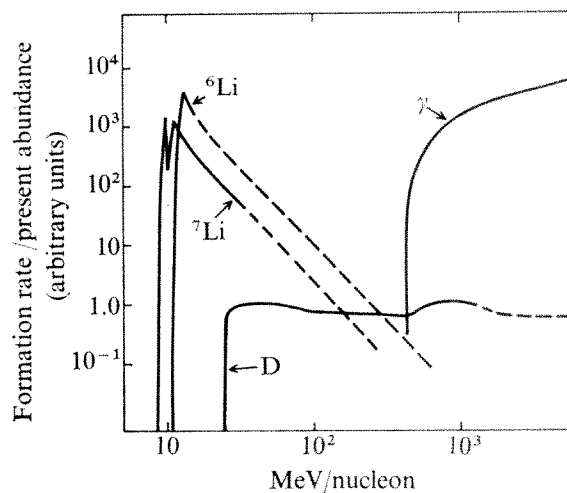
This constraint is clearly shown in Fig. 1. Particle fluxes with energies $\lesssim$ 300 MeV per nucleus generate very large Y_{Li}/Y_D ratios, and particles with energies > 500 MeV per nucleon produce high Y_γ/Y_D ratios. Models for a spallation origin of deuterium must rely on a flux of low energy particles and some mechanism for preferentially destroying the lithium or on a flux of high energy particles and some means of absorbing the γ rays photons. Particles with a narrow spread of energy ~ 400 MeV per nucleon could possibly produce reasonable Y_{Li}/Y_D and Y_γ/Y_D ratios, but models which require nearly monoenergetic particles are rightly considered *ad hoc* and are hard to justify on physical grounds.

It has been pointed out that the γ rays which form from the decay of π^0 can cause the photodisintegration of ^{4}He and thereby produce deuterium[20]. However, the γ rays preferentially produce electron–positron pairs[21,22] so that only a few percent of these photons can ultimately produce deuterium. Since $Y_\gamma/Y_D \sim 10^{-3}$ at present, any model which invokes photo-disintegration to form deuterium must also contain some mechanism for γ-ray destruction. Furthermore, if the π^0 are formed by high energy nuclear collisions, most of the deuterium is actually formed directly by $p+\alpha$ collisions, so that the photodisintegration reaction is of only minor importance.

Finally, it should be re-emphasised that, in the above discussion, the reactions on nuclei heavier than ^{4}He have been omitted. The inclusion of even a small amount of C, N or O nuclei would make it much more difficult to formulate an acceptable model for deuterium formation by spallation. A heavy element mass fraction ~ $Z = 10^{-4}$ leads to over-abundances of Li, Be and B relative to deuterium. This effectively eliminates such events as present-day solar flares[32].

Pregalactic cosmic rays

The high energy particles in the Galaxy could not have produced the galactic abundance of deuterium. The total number of these particles which have existed throughout the age of the Galaxy is far below what is required to produce significant amounts of deuterium, and one is still faced with the difficulty that lithium and other light nuclei would be overproduced. Rather than deal with the restrictive problem of whether deuterium could be formed in any particular type of event, consider the more general question: can one formulate any physically self-consistent model for a cosmic-ray origin of deuterium? The simplest possible model involves 'pregalactic cosmic rays' (only protons and α particles) of arbitrary energy interacting with an ambient medium of hydrogen and helium. A study of this sort has recently been undertaken by Epstein (unpublished) where particle spectra of the form

$$F(E) \propto (E+E_B)^{-2\cdot5} \qquad (2)$$

were used, with E_B an adjustable parameter. This spectral form has the advantage of being a simple one-parameter function which reduces to approximately the galactic cosmic-ray spectrum for suitable values of E_B ($\lesssim 1$ GeV per nucleon). Essentially similar results are obtained by using power law spectra of various indices and with high and/or low energy cutoffs.

When these cosmic rays interact with the ambient medium, deuterons, lithium nuclei, γ rays and other particles are formed. The initial Y_D/Y_{6Li} production ratio is shown in Fig. 2 as a function of E_B (this ratio is used because it provides a stronger constraint than Y_D/Y_{7Li}). Only for $E_B \gtrsim 20$ GeV per nucleon is the initial value of Y_D/Y_{6Li} comparable to the observed ratio. As the system evolves, the proton and α particle spectra are degraded by ionisation losses and nuclear collisions, the deuterons and lithium nuclei which were produced at high energies are slowed down or destroyed, and some of the γ rays are absorbed by the ambient medium. When all relevant production, destruction and energy loss processes are included the Y_D/Y_{6Li} ratio evolves, finally reaching the values indicated
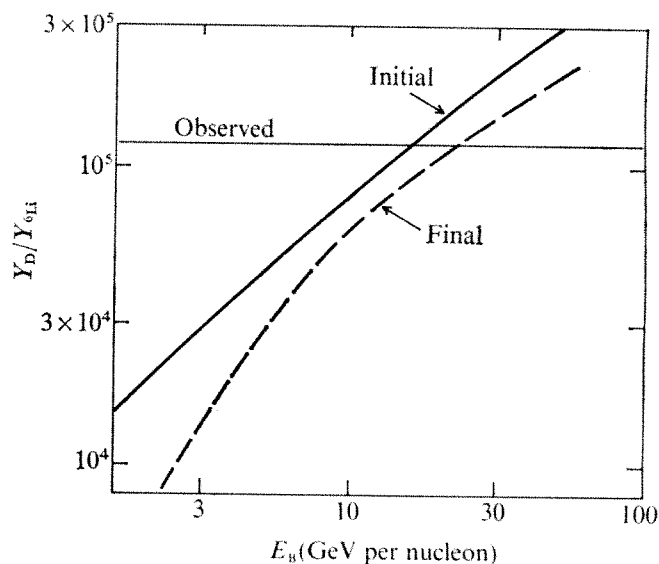
Fig. 2 Ratio of the abundances of deuterium and ^{6}Li which are produced by spallation reactions. The solid line is for spectra defined by equation (2). The dashed line is for spectra which are initially given by equation (2) and subsequently evolve. The observed abundance ratio is indicated.

by the dashed line in Fig. 2. Only for values of $E_B \gtrsim 30$ GeV per nucleon is the final Y_D/Y_{6Li} ratio acceptable. For these very hard cosmic ray spectra γ rays are, however, overproduced by a factor of > 10^4 relative to the diffuse background. Only if the source of γ rays were shielded by ~ 500 g cm^{-2} of matter would the observable flux of γ-ray photons be consistent with existing measurements. The energy which is required to give a deuterium abundance of $X_D = 2 \times 10^{-5}$ is 1.7×10^{18} erg g^{-1}, or 5×10^{62} erg for a $10^{11} M_\odot$ galaxy.

A spallation origin of deuterium cannot be totally dismissed on the basis of the above discussion, but can the necessary conditions for this mode of deuterium formation be realised in any reasonable astronomical event? The intense flux of cosmic rays, the low heavy element abundance and the large column density are certainly inconsistent with what is known about normal galaxies. Quasars, on the other hand, emit roughly the correct amount of energy, much of which is presumably derived from non-thermal particles, and they could provide sufficient shielding (a $10^8 M_\odot$ object with a 6-pc diameter has a column density > 10^3 g cm^{-2}). Observed quasars contain elements heavier than ^{4}He, but an earlier generation of 'extreme population II' quasars could be a possible site for deuterium production. The absorption of γ rays could also be achieved by the general ambient medium of the Universe if deuterium formation occurred at a sufficiently early epoch. This would require that deuterium production occur at redshifts > $z = 300$, shortly after recombination of the ionised matter in the Universe[23]. One can conjecture that an intense flux of high energy particles could be generated by an early generation of collapsing objects[24] or by energetic turbulence which results from the readjustment of the cosmic gas after its decoupling from the radiation field. Without other supportive evidence for such mechanisms, however, their invocation seems, again, quite *ad hoc*.

Shock waves

It has been suggested by Colgate[8,9] and by Hoyle and Fowler[10] that strong shock waves propagating through a low density gas could produce deuterium. The basic idea is that as a shock wave propagates through a low density medium, such as an extensive stellar envelope, it can exhibit a brief non-equilibrium phase in which high energy ion reactions ($\gtrsim 10$ MeV per nucleon) can occur; this 'ion precursor' phase arises partly from the initial collisions between the outwardly streaming ions and the

stationary low density matter, and partly from the relatively long time (compared to a nuclear reaction time) that is required for the shocked gas to relax to thermal equilibrium. As equilibrium is attained through electron heating and photon emission processes, most of the internal energy of the shocked material is transferred to the radiation field and the ions are quickly cooled. During the ion precursor, spallation reactions of protons on helium produce deuterium and free nucleons while spallation reactions on the CNO nuclei and α–α reactions produce Li, Be, and B. In the cool gas behind the precursor the free nucleons recombine to augment the deuterium abundance. Though this mechanism is an ingenious blend of the two basic deuterium production processes it runs into difficulties on several accounts.

First, the very existence of strong ion precursors is in doubt. Weaver[14] has been able to obtain solutions only for the shock wave structures which do not exhibit high temperature precursors. Furthermore, no one has been able to show that shock waves with ion precursors are actually physically self-consistent.

The second problem is that shock-wave nucleosynthesis tends to overproduce some of the light elements relative to deuterium[11,12]. Shock waves of energies up to 50 MeV per nucleon produce deuterium and other nuclei by direct spallation so that ^{7}Li is always overproduced. Stronger shock waves, however, spall nearly all the nuclei to free nucleons, thereby destroying the ^{7}Li overabundance and allowing for the subsequent synthesis of deuterium by ^{1}H(p,n)D. The unresolved question is whether an appreciable amount of matter can be processed by very strong shock waves without an even larger amount of matter being exposed to weaker shock waves which would preferentially produce ^{7}Li. Colgate[25] contends that at low shock velocities the diffusion of radiation damps the development of a precursor (and thus prevents ^{7}Li from forming), and that at high shock velocities the radiation does not keep pace with the shock wave so that the precursor develops and deuterium forms. As mentioned before, the detailed calculations of Weaver do not support this picture.

If the arguments in favour of deuterium production in shock waves were accepted there would still remain the question of what the origin of these shock waves could be. Normal supernovae appear to be highly unlikely sites. Using the most favourable assumptions (low CNO abundances, efficient acceleration of the shock waves) Epstein *et al.*[12] found that energies $> 10^{53}$ erg would have to be released in a typical $5M_\odot$ type II supernova; this is far greater than any observational inferences[15]. Pregalactic supermassive objects could have, perhaps, undergo extremely violent explosions, and few constraints can be put on their energetics. Nevertheless, there is no reason to expect that they could be any better suited for producing deuterium than current supernovae.

Hot explosions

Deuterium can be formed if matter is brought to high temperatures where heavy nuclei decompose to free nucleons ($T > 10^9$–10^{10} K, depending on density) and then rapidly expanded and cooled. The deuterium is formed by ^{1}H(n,γ)D reactions, and it is also destroyed by thermonuclear processes, the most important destruction reactions being (D(d,n)^{3}He and D(d,p)^{3}H as long as $Y_D \gtrsim 10^{-5}$. What fraction of the deuterium which is formed in an explosive event actually escapes destruction? If the densities of neutrons and protons are nearly equal there are two possible sequences by which deuterium can be formed and also survive (the case in which there is a large excess of neutrons is considered later):

(1) The expansion may be rapid enough so that some neutrons and protons do not recombine until the gas cools to the 'freeze-out' temperature ($T \lesssim 3 \times 10^7$ K), when the charged-particle reactions are slower than the expansion. The deuterium which is subsequently formed is ejected without much thermonuclear destruction.

(2) Deuterium forms while the gas is still hot but the expansion

is rapid enough so that deuterium is ejected without being completely destroyed.

Neutrons and protons are consumed at a rate

$$\frac{dY_n}{dt} = \frac{dY_p}{dt} = -4 \times 10^4 \rho \, Y_n Y_p \, \text{s}^{-1} \qquad (3)$$

where ρ is in g cm^{-3}. Define $\rho_0 \tau_{exp} = \int \rho dt$ where the integration is from the time at which the nucleons start being consumed (when $T \gtrsim 3 \times 10^9$ K) until freeze out. The upper limit on the number of free neutron–proton pairs at freeze out and hence on the abundance of deuterium which forms after freeze out is

$$Y_D \lesssim (4 \times 10^4 \rho_0 \tau_{exp})^{-1} \qquad (4)$$

Deuterium which forms at high temperatures is subject to destruction by thermonuclear reactions. The rate at which deuterium is burnt by D + D reactions at $T \sim 10^9$ K is[26]

$$\frac{dY_D}{dt} \simeq -10^7 \rho Y_D^2 \qquad (5)$$

so that the amount of deuterium which forms at high temperatures and survives until freeze out is limited by $Y_D \lesssim (10^7 \rho_0 \tau_{exp})^{-1}$. Since this is far less than the maximum amount of deuterium production after freeze out, equation (4) represents a strong upper limit to the net deuterium production in a hot explosion.

Since only a small fraction of the matter in the Universe could have been raised to high temperatures and then rapidly expanded (except, of course, in the big bang), the deuterium fraction in the ejecta has to be greater than the interstellar value of $\sim 10^{-5}$. Taking the value to be $Y_D > 10^{-3}$ and using equation (4) gives

$$\rho_0 \tau_{exp} < 2 \times 10^{-2} \text{ g s cm}^{-3} \qquad (6)$$

Can such rapid expansion rates be attained in realistic explosive events? To obtain a lower limit to τ_{exp}, consider an homogeneous sphere of mass M and density ρ expanding so that its surface moves with some speed less than the velocity of light. Taking $\tau_{exp} = \rho/(d\rho/dt)$ gives

$$\rho \tau_{exp} \gtrsim 0.9 \, \rho^{2/3} \, (M/M_\odot)^{1/3} \text{ g s cm}^{-3} \qquad (7)$$

The density of matter at 10^9 K has to be $\gtrsim 8$ g cm^{-3} for the energy density of the radiation $> \rho c^2$. (Only in the big bang can this condition be reasonably violated.) Expansion rates which are consistent with equation (6) thus require the mass of the exploding region to be $\lesssim 10^{-7} M_\odot$. Since this is much too small to correspond to any reasonable astronomical event, we conclude that significant amounts of deuterium cannot survive a hot explosion.

Disrupted neutron stars

Can neutron stars provide the free neutrons which are needed for deuterium synthesis? Lattimer and Schramm (ref. 16, and unpublished) investigated a mechanism whereby a neutron star falling towards a black hole (for example, a binary system decaying by gravitational radiation) is tidally disrupted and ejects some neutron-rich matter from the system (the 'tube-of-toothpaste' effect), and others (refs 30, 31 and A. G. W. Cameron and J. Truran, unpublished) have shown how neutron-rich matter can be ejected by MHD instabilities from the collapsing core (a 'hot neutron star') of a highly evolved star. For deuterium to be produced in this type of process the neutron-rich matter has to interact with protons before (1) all the free neutrons are consumed in forming heavy nuclei or (2) before the neutrons decay to protons. Following Lattimer *et al.*[17] it seems that, as the neutron-star matter expands, the few protons which are initially present ($Y_p \simeq 0.03$–0.04) clump

together with the neutrons to form clusters. The excess neutrons 'boil off' until the nuclear clusters reach the neutron drip line. As the density decreases further, β decays begin to occur in the nuclei; the increase in proton number permits the recapture of neutrons. Nuclei up to mass $A \sim 300$ are formed and then fission produces more seed nuclei. Neutrons are consumed as rapidly as permitted by the β decays and fissions of the heavy neutron-rich nuclei. The timescale for this to occur is ~ 0.1 s. (Similar processes occur for the hot neutron star except the initial Y_p are 2 or 3 times larger and the neutron depletion time scales could be considerably longer.) Expansion rates more rapid than this would allow some free neutrons to escape.

Even if some free neutrons do escape, however, they decay in $\sim 10^3$ s and do not produce any deuterium unless they interact with hydrogen-rich matter with a density of at least 10^{-8} g cm^{-3}. Densities of this magnitude are associated with stars or stellar envelopes. Models of deuterium formation which require a large number of neutron stars to interact with black holes and to spew neutron-rich matter into nearby stars are probably beyond the range of reasonable speculation. Hot neutron stars, however, are still surrounded by the pre-supernova stellar envelopes. For the hot neutron-rich material to be able to interact with the overlying hydrogen mantle several conditions must be satisfied: (1) the β-decay rates for very neutron-rich nuclei must be ~ 1 s; (2) the neutrons cannot interact with the heavy elements which surround the core, and (3) the neutron-rich matter has to be ejected rapidly enough so that it can reach the hydrogen mantle (a distance of 10^{10}–10^{12} cm from the core) in $< 10^3$ s. Any or all of these conditions are likely to be violated, but for now this remains a viable model.

Conclusion

Big-bang nucleosynthesis provides an elegant theory for the synthesis of deuterium and helium. It requires only the simplest cosmological assumptions and is consistent with all well established astronomical data. Contrasted to this, post-big-bang deuterium production requires extremely violent and exotic events, the existence of which is certainly doubtful. From the perspective of our current understanding of the chemical evolution of the Universe, big-bang nucleosynthesis is by far the most reasonable site for the origin of deuterium. What are the prospects that this conclusion may be altered or made more firm? Two approaches may be fruitful. First, one can search for direct evidence of post-big-bang deuterium formation such as inhomogeneities in the abundance of deuterium or enhancements in the abundances of lithium, beryllium, boron and γ rays, which may be associated with deuterium production. Variations of this sort either in our Galaxy (ref. 27, and A. A. Penzias, P. G. Wannier, R. W. Wilson and R. A. Linke, unpublished) or in extragalactic objects such as quasars (T. Adams, unpublished) would be of great importance. The results to date

indicate a possible decrease in the deuterium abundance near the galactic centre (A. A. Penzias, P. G. Wannier, R. W. Wilson and R. A. Linke, unpublished); this is consistent with any pregalactic origin of deuterium followed by destruction of deuterium by stellar processing. The second approach, which is being vigorously pursued by a number of researchers (see ref. 5 and references therein), is to determine the local mean mass density of the Universe. Considerations of the dynamics of galaxies and groups of galaxies has led to the tentative result that the mean density is at least 5×10^{-31} g cm^{-3}. If further studies could raise this limit by even a firm factor of two or three (still leaving ρ below the critical density ρ_c), then the standard cosmological picture for the origin of deuterium would have to be seriously questioned and exotic pregalactic or hot neutron star models might have to be invoked.

It may turn out that deuterium was formed during the initial hot phase of the big bang or during subsequent epochs. In either case the abundance of this isotope provides an intriguing clue to evolution of the early Universe.

1 Wagoner, R. V., *Astrophys. J.*, 179, 343 (1973).
2 Rogerson, J. B., and York, D. G., *Astrophys. J. Lett.*, 106, L95 (1973).
3 Pasachoff, J. M., and Cesarsky, D. A., *Astrophys. J.*, 193, 65 (1974).
4 Schramm, D. N., and Wagoner, R. V., *Phys. Today*, 27, 12, 40 (1974).
5 Gott, J. R., Gunn, J., Schramm, D. N., and Tinsley, B., *Astrophys. J.*, 194, 543 (1974).
6 Epstein, R. I., and Petrosian, V., *Astrophys. J.*, 197, 281 (1975).
7 Clayton, D. D., *Principles of Stellar Evolution and Nucleosynthesis* (McGraw-Hill, New York, 1968).
8 Colgate, S. A., in *Explosive Nucleosynthesis* (edit. by Schramm, D. N., and Arnett, W. D.), (University of Texas Press, Austin, 1973).
9 Colgate, S. A., *Astrophys. J.*, 187, 321 (1974).
10 Hoyle, F., and Fowler, W., *Nature*, 241, 384 (1973).
11 Epstein, R. I., Arnett, W. D., and Schramm, D. N., *Astrophys. J. Lett.*, 190, L13 (1974).
12 Epstein, R. I., Arnett, W. D., and Schramm, D. N., *Astrophys. J. Suppl.*, 31, 111 (1976).
13 Weaver, T. A., and Chapline, G. F., *Astrophys. J. Lett.*, 192, L57 (1974).
14 Weaver, T. A., *Astrophys. J.* (in the press).
15 Reeves, H., *A. Rev. Astr. Astrophys.*, 12, 437 (1974).
16 Lattimer, J. M., and Schramm, D. N., *Astrophys. J. Lett.*, 192, L145 (1974).
17 Lattimer, J. M., Mackie, F. D., Ravenhall, D. G., and Schramm, D. N., *Bull. Am. astr. Soc.*, 7, 546 (1975).
18 Cameron, A. G. W., *Space Sci. Rev.*, 15, 121 (1973).
19 Stecker, F. W., in *Gamma-Ray Astrophysics* (edit. by Stecker, F. W., and Trombka, J. I.), (NASA SP-339, Washington, D.C., 1973).
20 Melik-Alaverdyan, Yu. K., *Astrofizika*, 10, 123 (1974).
21 Stecker, F. W., *Cosmic Gamma Rays* (NASA SP-249; Washington, DC, 1971).
22 Arkatov, Yu. M., et al., *Sov. J. Nucl. Phys.*, 19, 598 (1975).
23 Harrison, E. R., *A. Rev. Astr. Astrophys.*, 11, 155 (1973).
24 Peebles, P. J. E., and Dicke, R. H., *Astrophys. J.*, 154, 891 (1968).
25 Colgate, S. A., *Astrophys. J.*, 195, 493 (1975).
26 Fowler, W. A., Caughlan, G. R., and Zimmerman, B. A., *A. Rev. Astr. Astrophys.*, 5, 525 (1967).
27 Ostriker, J. P., and Tinsley, B. M., *Astrophys. J. Lett.* (in the press).
28 Bodansky, D., Jacobs, W. W., and Oberg, D. L., *Astrophys. J.*, 202, 222 (1975).
29 Ramaty, R., and Lingenfelter, R. E., *Astrophys. J.*, 155, 587 (1969).
30 Meier, D., Epstein, R. I., Schramm, D. N., and Arnett, W. D., *Astrophys. J.*, 204, 869 (1976).
31 LeBlanc, J. M., and Wilson, J. R., *Astrophys. J.*, 161, 541 (1970).
32 Coleman, G.D., and Worden, S.P., *Astrophys. J.*, 205, 475 (1976).

Orthogonal mode emission in geometric models of pulsar polarisation

D. C. Backer
Radio Astronomy Laboratory, University of California, Berkeley, California 94720

John M. Rankin
Astronomy Department, Cornell University, Ithaca, New York 14853

D. B. Campbell
Arecibo Observatory, Arecibo, Puerto Rico 00612

The varying incidence of orthogonal 'states' of linear polarisation across the pulse of PSR 2016 + 28 is discussed in terms of a model with subpulse beams drifting round the pole of a rotating magnetic dipole.

THE position angle associated with the average linearly polarised radiation from many pulsars varies across the pulse in a manner consistent with a simple geometrical model first given by Radhakrishnan and Cooke[1,2]. The model postulates polarised radiofrequency emission having properties character-
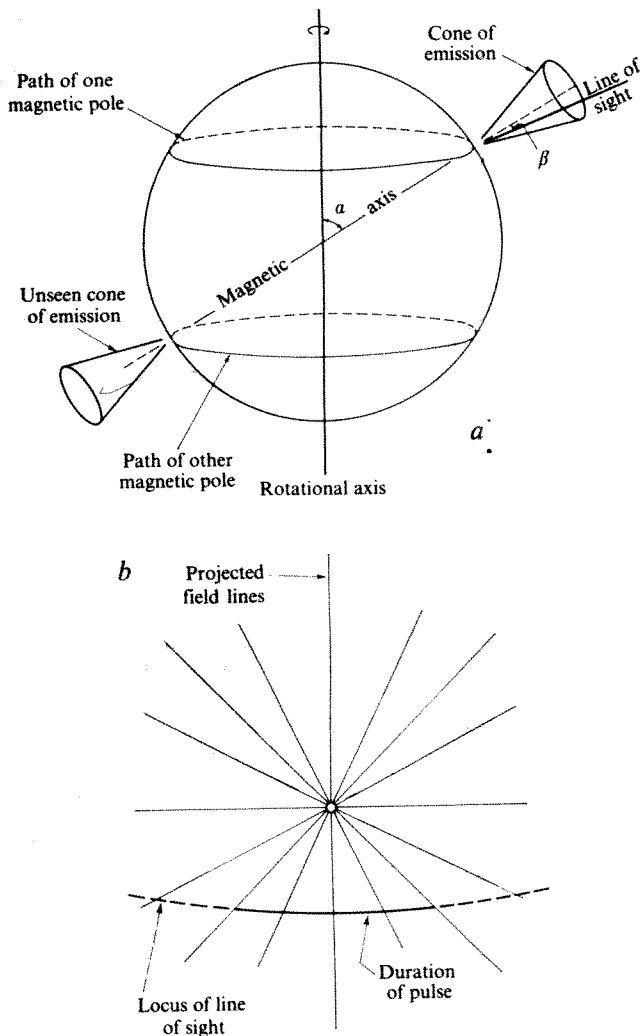
Fig. 1 *a*, Schematic representation of the Radhakrishnan and Cooke pulsar model[1]. Emission is associated with the polar regions of an oblique, rotating, magnetic dipole. *b*, Field geometry near the pole of the star projected on to the celestial sphere. The radiofrequency radiation is expected to have polarisation characteristic of curvature radiation from particles accelerated along the field lines[1]. *c*, Model waveforms of linear polarisation angle for a magnetic dipole axis inclined at 60° to the rotation axis; curves for observers at various polar angles are indicated. *d*, 430-MHz Arecibo observations of two pulsars (PSR 0525+21, left; PSR 1915+13, right) with linear polarisation angle morphologies which conform well with the Radhakrishnan–Cooke model angle waveforms in (*c*). The position angles are not absolute.
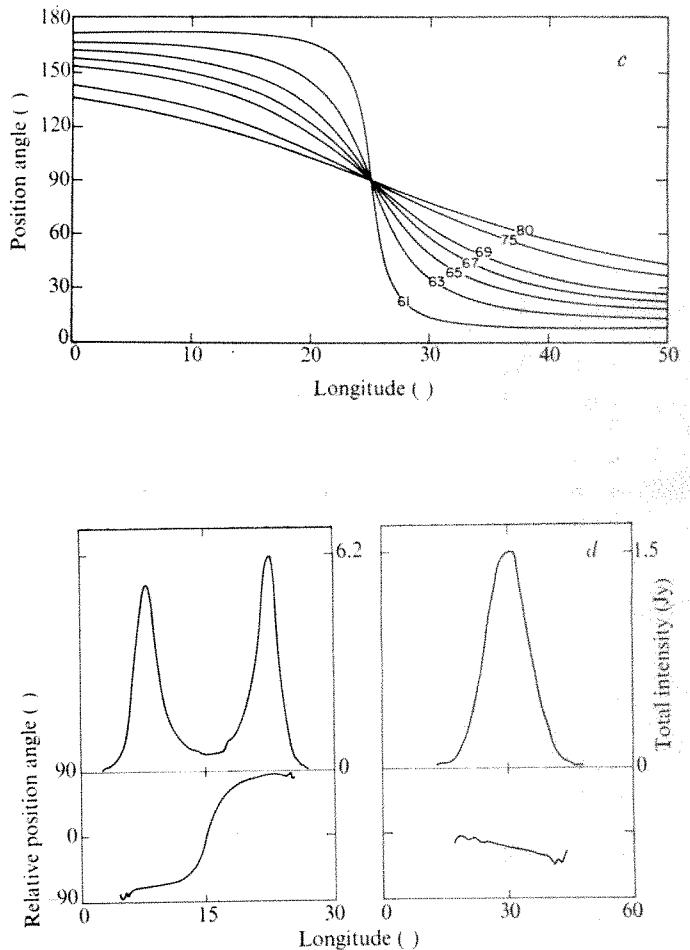
istic of the radiation from relativistic particles accelerated along diverging field lines near the poles of an oblique, rotating, magnetic dipole. Figure 1*a* shows the general configuration of the two emitting regions and Fig. 1*b* gives a view of the field lines as projected on to the celestial sphere. Position angles associated with the linearly polarised component of the radiation in this model are expected to coincide with the magnetic field projections in Fig. 1*b*. Thus the position angles are expected to progress monotonically through some portion of 180° during the pulse as the rotation of the star carries the polar cone of emission across the line of sight. Figure 1*c* gives model angle waveforms calculated for a dipole inclined at 60° and a range of observer inclinations; Fig. 1*d* presents observations of two objects as examples of the many pulsars which are nominally consistent with the model (see also, for instance, ref. 3).

There is, however, a second class of pulsars which displays more complicated polarisation angle behaviour that cannot be adequately accounted for on the basis of the above geometric model. In particular, there are several objects in the Arecibo pulsar polarisation survey[4] which exhibit a rapid swing in the average polarisation angle by ~ 90° (for example, pulsars 1541+09, 1604−00, 1944+17, 2016+28 and 2020+28, see Fig. 2). Clearly the Radhakrishnan and Cooke model can account for only a single rapid change of ~ 180°. Furthermore,

Fig. 2 430-MHz Arecibo observations of two pulsars which exhibit complex polarisation angle morphology (PSR 1944+17, left; PSR 2016+28, right). Rapid transitions of ~ 90° are evident in both cases. The position angles are not absolute.
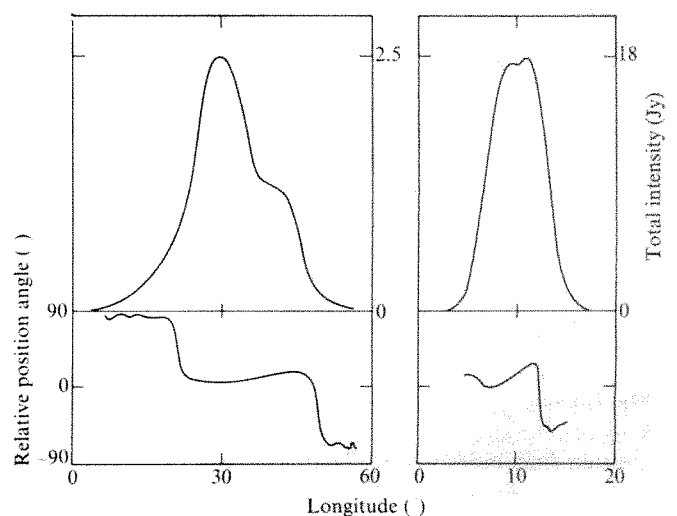
Long. (degrees)	$\langle I \rangle$ Jy	$A1$ (degrees)	$A2$	Polarisation angle (degrees) 0 — 180	
0.0	2.6	41	54	|.211.1 /1 . 1111. 1 |	•
0.4	3.6	34	49	|.22111.. ...1.... • ...|	•
0.8	5.1	29	44	|11221................. • ..|	1
1.2	7.3	28	48	|.1221 1. ....|	1
1.5	10.4	28	41	|.1221... ....1.. • ..|	2
1.9	14.4	31	47	|.1221... ....11.. ..|	3
2.3	19.7	31	48	|.1121 1........11... • ...|	4
2.7	25.8	31	51	|.1121 1.... ..11.... ...|	4
3.1	33.2	32	55	|.1111 11........111.. ....|	5
3.5	41.0	32	59	|.1111 1........111... .....|	6
3.9	49.7	29	62	|..111 1........111... ....|	7
4.3	59.3	26	61	|..111 1.........11... ...|	7
4.6	67.5	25	61	|..11 21..... ..111. .. ...|	7
5.0	75.4	25	58	|..11 21.........111... ...|	7
5.4	81.9	25	56	|..12 211.........11.. ...|	7
5.8	85.3	26	55	|..12 211.............1.. • ....|	8
6.2	87.5	27	54	|..11 211.... ...111. • ...|	8
6.6	87.9	28	53	|..11 221.............11.. ..|	8
7.0	87.7	30	55	|..11 221.............11... • • |	8
7.4	88.8	32	56	|...12 21..... .....1..... ..|	8
7.7	89.3	33	56	| ..12 211.... ...1... ..|	7
8.1	88.6	35	57	|...11 221.... ....1.... • |	8
8.5	89.2	36	61	|....12 21.............1......|	8
8.9	90.3	39	65	|....12 21.............11..... |	8
9.3	90.6	41	70	| ....12 11.............11.. • |	8
9.7	90.1	43	76	| ....12 11.............11..... |	8
10.1	88.3	43	84	| ....11 11..........121.... |	8
10.5	85.7	43	91	| ....11 11..........121.... |	8
10.8	81.9	39	95	| ....11 11... .....122.... |	8
11.2	77.0	27	100	| ..11 11.... ...221... |	8
11.6	70.6	178	107	| ....111..... ...121... |	8
12.0	62.9	169	110	|.......111..... ....132.. |	8
12.4	54.9	173	113	| ......11.... • • ..1221..|	7
12.8	44.9	166	114	| ......11.... • ...1221.. |	7
13.2	34.8	161	117	|.......11..... ..1221.. |	6
13.6	26.1	156	123	|. ....1.... .. ..1221. |	5
13.9	19.6	156	128	| .. ....1... ..1321..|	4
14.3	14.9	156	134	| ...1... • ..12311.|	4
14.7	11.7	158	136	| • ....1.. ..1232..|	3
15.1	9.3	158	136	|.... ...... • ....1232..|	2
15.5	7.6	160	138	|... .... • ...11321.|	2
15.9	6.2	161	133	|.... .... .11321.|	1
16.3	5.2	165	126	|.1 • ...1.. ....1211.|	1
16.6	4.4	165	124	|.1.. .... • ..1221.|	1
17.0	3.7	166	121	|1.. ... .1. • .....1221.|	1
17.4	3.0	163	117	|..1. • 11. ... • .1121.|	•
17.8	2.4	164	131	|. • 1.. • 1 ...1221|	•

Fig. 3 Probability distribution of linear polarisation angle at a sequence of longitudes across the pulse of pulsar 2016+28. Long. is the relative longitude in degrees, $\langle I \rangle$ is the amplitude of the total intensity average waveform in Jy, $A1$ is the vector-average position angle waveform in degrees (also plotted superposed on the distributions for comparison), and $A2$ is the corresponding arithmetic (amplitude independent)-average polarisation angle waveform. The display consists of probability densities of polarisation angle (in units of 10%) (with a decimal point indicating 0.05–5.0%) per resolution cell. The distributions reflect only those data which exceeded an amplitude threshold computed from the off-pulse noise baseline; the final column on the right gives the fraction of data above the noise threshold (in units of 10%).

90° rotation has been found to occur within subpulses of PSR 0809+74 (ref. 5), from subpulse to subpulse in PSR 2303+30 (ref. 4) (two objects dominated by drifting subpulses), and has been noted in several contexts recently[6].

Preferred polarisation angles

Statistical studies of the Arecibo polarisation survey data[7] indicate that most objects have, at a given longitude (360° longitude = 1 pulse period), two, nearly orthogonal, preferred polarisation angles of emission, and also that the relative proportion of the two '90° states' often changes markedly across the pulse waveform (see also ref. 6). The varying proportion of the orthogonally-polarised constituents serves to distort a 'fundamental' angle waveform which, in most cases, is found to be consistent with the Radhakrishnan and Cooke model. (In other words, angle waveforms computed by averaging the data |90°| (instead of |180°|) would compare well (|90°|) with those in Fig. 1c.) Specifically, the characteristic bimodal nature of the position angle distribution arises, at least in part, from the rapid transition between states within
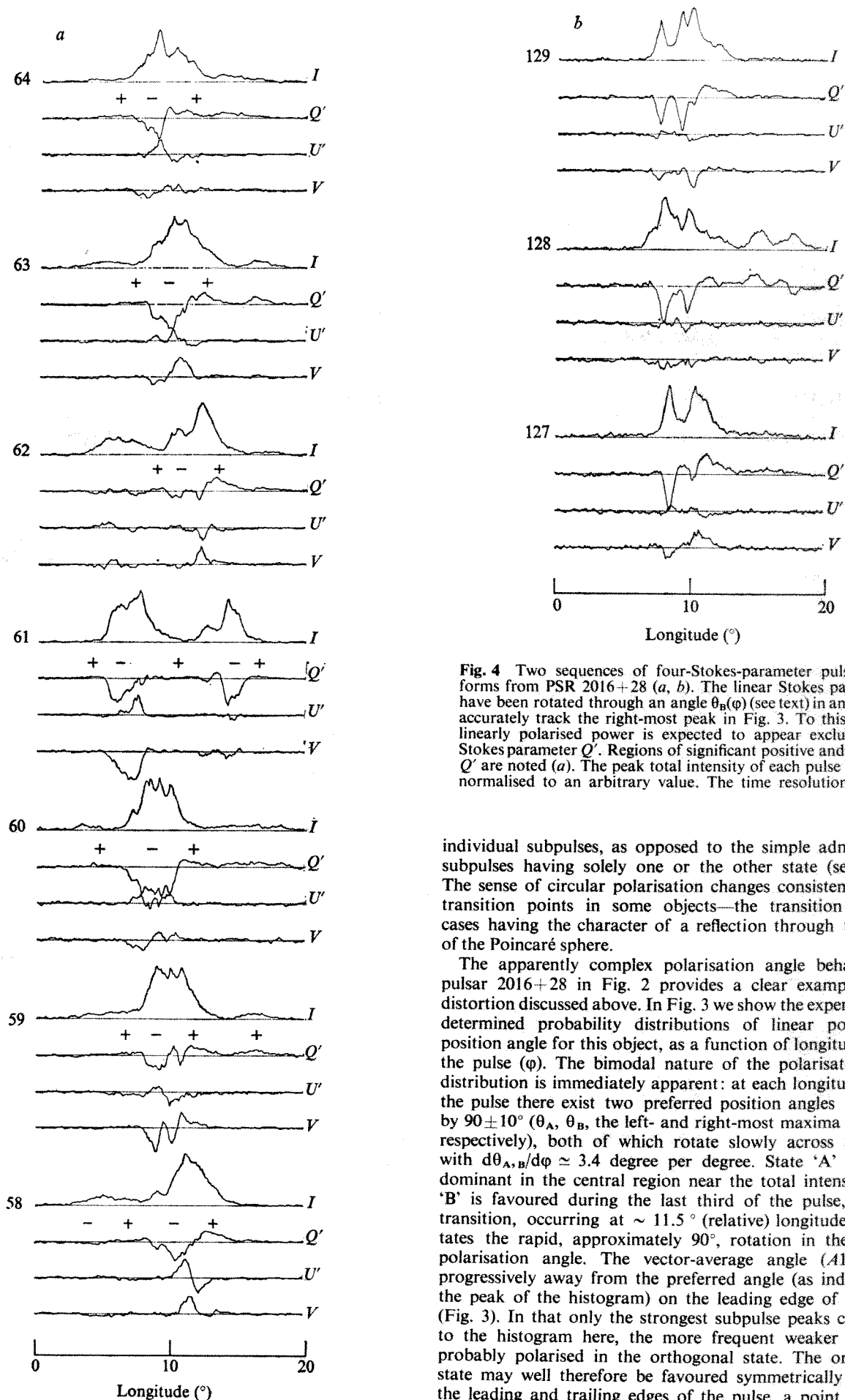
Fig. 4 Two sequences of four-Stokes-parameter pulse waveforms from PSR 2016+28 (*a*, *b*). The linear Stokes parameters have been rotated through an angle $\theta_B(\varphi)$ (see text) in an effort to accurately track the right-most peak in Fig. 3. To this end the linearly polarised power is expected to appear exclusively in Stokes parameter Q'. Regions of significant positive and negative Q' are noted (*a*). The peak total intensity of each pulse has been normalised to an arbitrary value. The time resolution is 0.5°.

individual subpulses, as opposed to the simple admixture of subpulses having solely one or the other state (see ref. 6). The sense of circular polarisation changes consistently at the transition points in some objects—the transition in these cases having the character of a reflection through the origin of the Poincaré sphere.

The apparently complex polarisation angle behaviour of pulsar 2016+28 in Fig. 2 provides a clear example of the distortion discussed above. In Fig. 3 we show the experimentally determined probability distributions of linear polarisation position angle for this object, as a function of longitude within the pulse (φ). The bimodal nature of the polarisation angle distribution is immediately apparent: at each longitude within the pulse there exist two preferred position angles separated by $90 \pm 10°$ (θ_A, θ_B, the left- and right-most maxima in Fig. 3, respectively), both of which rotate slowly across the pulse with $d\theta_{A,B}/d\varphi \simeq 3.4$ degree per degree. State 'A' is clearly dominant in the central region near the total intensity peak, 'B' is favoured during the last third of the pulse, and the transition, occurring at $\sim 11.5°$ (relative) longitude, precipitates the rapid, approximately 90°, rotation in the average polarisation angle. The vector-average angle ($A1$) moves progressively away from the preferred angle (as indicated by the peak of the histogram) on the leading edge of the pulse (Fig. 3). In that only the strongest subpulse peaks contribute to the histogram here, the more frequent weaker data are probably polarised in the orthogonal state. The orthogonal state may well therefore be favoured symmetrically on both the leading and trailing edges of the pulse, a point which is discussed further below.

Subpulse polarisation

Having investigated the occurrence of the two orthogonally-polarised angle states in the context of the average polarisation angle waveform, it is now clearly of interest to explore the nature of subpulse polarisation in the same light. Since the linear polarisation at any point in the pulse is observed to favour only two orthogonal position angles, the linear Stokes parameters at each point can be defined so that one linear component, say U', is always zero and the other, Q', reverses sign between the two states. This arrangement of the linear Stokes parameters (Q', U') is effected for all longitudes φ by rotating the reference angle from position angle zero to θ_B (φ). Two sequences of pulses from PSR 2016+28, in all four Stokes parameters, I, Q', U' and V, are shown in Fig. 4. From Fig. 4 it can be seen that, although the Stokes parameter U' is imperfectly nulled, the linearly polarised power preponderates in Q' and so our general view seems valid.

Models

We wish to interpret the Q' morphology in terms of the properties of subpulse beams according to the general tenets of the model proposed by Radhakrishnan and Cooke, whose views have been both extended and given a firm theoretical foundation by Sturrock[8] and by Ruderman and Sutherland[9]. In addition, Ruderman[10] explored the possibility that the drifting subpulse phenomenon is produced by narrow subpulse beams circulating round the dipole axis. The dimensions and motions of such subpulse beams seem to be consistent with a hollow-cone average beam—a concept suggested by Komesaroff[2] which now seems to be in agreement with much experimental evidence[11]. Figure 5 summarises the various relevant features of these models in a single drawing where we have yet to speculate on the polarisation structure of the subpulse beams (indicated by $+/-$ signs). Finally, the presence of two unresolved components in the total intensity average waveform of 2016+28 (evident in Fig. 2 as well as in other observations[11]) suggests that our line of sight passes somewhere near the inside edge of the cone as indicated in Fig. 5.

The Q' Stokes parameter

The Q' waveforms in Fig. 4 provide a tracing of the two orthogonal modes of emission (the sense of Q' distinguishing be-tween the two) as the rotation of the star carries the pulsar's polar region across our line of sight. It is evident in Fig. 4a that central subpulses (those near the total intensity peak of the pulse) are most frequently characterised by a large negative excursion of Q', bounded on one or both sides by smaller positive regions. The configuration depicted in Fig. 5 (a, c, e)—that is, a total power beam with a central lobe of one emission state $(Q' < 0)$, bounded symmetrically by two sidelobes of the orthogonal state $(Q' > 0)$—seems to be the simplest time invariant beam largely compatible with the observed subpulse polarisation characteristics. In particular, the subpulse beams in Fig. 5 (h, i) do not reproduce the observations satisfactorily. Depending on the manner in which the line of sight cuts the subpulse beams, $a-e$, there is substantial flexibility in the kinds of subpulses produced: a well centred subpulse beam (Fig. 5c) produces a symmetrical, negative-Q'-dominated subpulse (Fig. 5C), whereas less central beams (Fig. 5a, e) result in increasingly asymmetrical, positive-Q'-dominated subpulses (Fig. 5A, B, D, E). Although it is difficult to identify positive-Q'-dominated subpulses at the extreme leading and trailing edges of the pulse (Fig. 4) because of the low intensity, Fig. 3 provides evidence that the orthogonal $(Q' > 0)$ state was favoured symmetrically at the extremes of the pulse. That the symmetry is not more pronounced seems to indicate that we are able to see subpulse circulation further in the direction of Fig. 5a than in that of 5e. Furthermore, the subpulses of PSR 2016+28 generally seem to consist of one or more micro-pulses[12] which appear to perturb the foregoing picture, often producing a narrower version of the subpulse polarisation beam apparently superimposed on the broader beam. Some examples of subpulses dominated by micropulses can be seen in Fig. 4b. The actual nature of the radiation is doubtlessly much more complicated than we have outlined here and observations with much finer time resolution are required to study polarisation properties at the micropulse level.

Discussion

Although it is not yet clear whether every aspect of the polarisation properties of PSR 2016+28 are typical of pulsars, some generalisations can be made from the observations. Statistical studies similar to Fig. 3 show both that the two orthogonal '90° states' of emission are a characteristic feature of pulsar
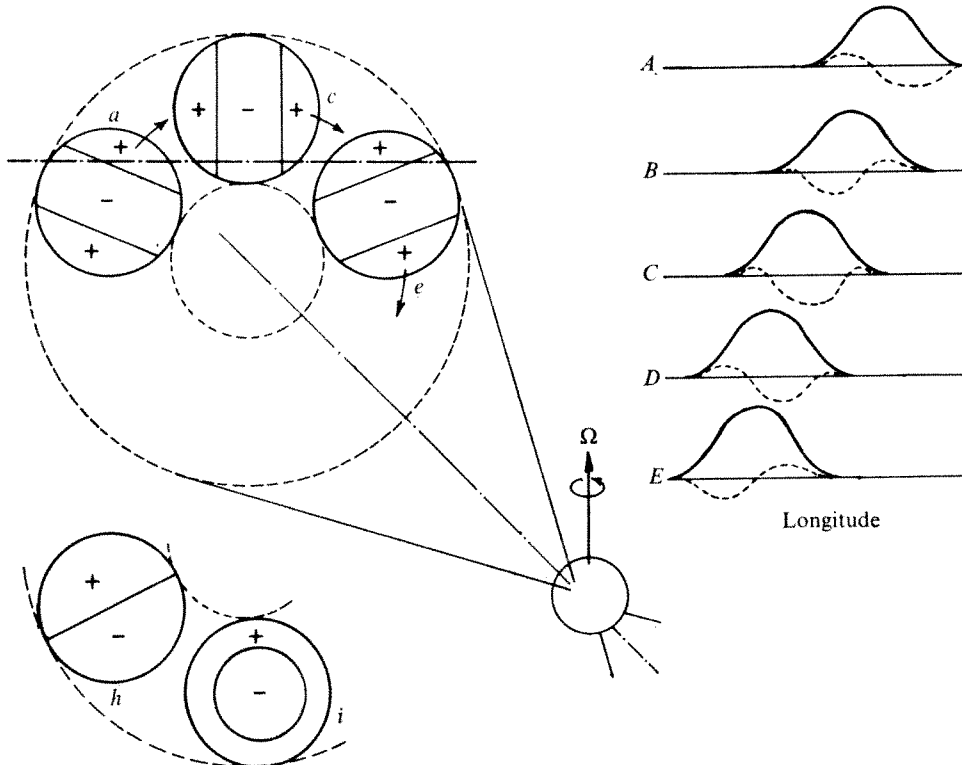
Fig. 5 A schematic drawing of the subpulse beam model for PSR 2016+28 under consideration (see text). This is essentially Ruderman's elaboration[10] on the Radhakrishnan and Cooke model[1], wherein subpulse beams are thought to circulate within a hollow-cone average beam, except that we now specifically consider the nature of subpulse polarisation ($+/-$ indicates the two orthogonal lobes of subpulse polarisation). The dashed curve traces the approximate path of the line of sight across the polar emission region. Sketches of subpulse waveforms corresponding to cuts through three of the illustrated subpulse beams (a, c, e) are given on the right (A, C, E), as well as two intermediate cases (B, D). Polarised subpulse lobe structures (h) and (i) do not seem to correspond to the observations.

Longitude

radiation and that a much larger group of pulsars can now be well described by a simple geometric model when the nature of their radiation is appropriately considered. The polarised individual pulse waveforms of PSR 2016+28, as well as other objects with strong drifting subpulse modulation, have a straightforward interpretation wherein subpulse beams with a rigid polarisation pattern are viewed in different orientations as time progresses. Pulsars 1919+21, 1944+17, 2016+28, the second component of 2020+28 and 2303+30 in the Arecibo survey as well as PSRs 0031−07 and 0809+74 in other studies[6] all bear out various features of this view. On the basis of current pulsar models we have suggested one method of producing subpulse beams which change their orientation on the sky as required: narrow subpulse beams with time-invariant polarisation patterns circulate around the dipole axis maintaining a fixed orientation to the plane of the magnetic field. Our speculations regarding the polarised lobe structure of the subpulse beams are less certain but may have some generality. Neverthe-

less, we hope that they will serve to prompt some further theoretical investigations of emission mechanisms and propagation effects[13] in pulsar magnetospheres.

Arecibo Observatory is operated under contract to the NSF by Cornell University.

1 Radhakrishnan, V., and Cooke, D. J., *Astrophys. Lett.*, 3, 225–229 (1969).
2 Komesaroff, M. M., *Nature*, 225, 612–614 (1970).
3 Morris, D., Schwartz, U. J., and Cooke, D. J., *Astrophys. Lett.*, 5, 181–186 (1970).
4 Rankin, J. M., Campbell, D. B., and Backer, D. C., *Astrophys. J.* 188, 690–613 (1974).
5 Taylor, J. H., Huguenin, G. R., Hirsch, R. M., and Manchester, R. N., *Astrophys. Lett.*, 9, 205–208 (1971).
6 Manchester, R. N., Taylor, J. H., and Huguenin, G. R., *Astrophys. J.*, 196, 83–102 (1975).
7 Backer, D. C., *Goddard Space Flight Center*, Rept X-693-76-43 (1976).
8 Sturrock, P. A., *Astrophys. J.*, 164, 529–556 (1971).
9 Ruderman, M., and Sutherland, P. G., *Astrophys. J.*, 196, 51–72 (1975).
10 Ruderman, M., *Astrophys. J.*, 203, 206–208 (1976).
11 Backer, D. C., *Astrophys. J.* (in the press).
12 Backer, D. C., *Astrophys. J.*, 182, 245–276 (1973).
13 Blandford, R. D., and Scharlemann, E. T., *Mon. Not. R. astr. Soc.*, 174, 59–86 (1976).

Structure of locust adipokinetic hormone, a neurohormone that regulates lipid utilisation during flight

Judith V. Stone & William Mordue

Department of Zoology

Karen E. Batley & Howard R. Morris

Department of Biochemistry, Imperial College, London SW7 2AZ, UK

Adipokinetic hormone, isolated from locust corpora cardiaca, has been identified as a blocked peptide: PCA–Leu–Asn–Phe–Thr–Pro–Asn–Trp–Gly–Thr–NH₂. The detailed structure is based on mass spectrometric data, substantiated in part by dansyl-Edman and carboxypeptidase data on thermolytic fragments. This is the first peptide hormone from an insect neuroendocrine organ to be fully characterised.

In insects, it has been shown (for reviews, see refs 1 and 2) that peptide and protein hormones are produced by neurosecretory cells or glandular cells associated with the neuroendocrine system. These neurosecretory cells are in the brain, the corpora cardiaca and the medial nervous system. Much work has been carried out investigating the factors present in the corpora cardiaca, which are neurohaemal organs that both store and release neurosecretory products synthesised in the brain, and contain intrinsic glandular cells producing their own secretions. So far, factors that affect cuticle tanning[3], water balance[4], carbohydrate metabolism[5−7], muscle contraction[8−10] and lipid metabolism[11−13] have all been found in the corpora cardiaca from various insects, and many of these factors have been shown to be of a peptide nature. Until now, however, none of these active substances has been purified and fully characterised. We report here the purification and identification of one of these factors, the adipokinetic hormone from the locust.

The existence of an adipokinetic hormone in locusts, located in the corpora cardiaca and concerned with flight activity, was first suggested by Mayer and Candy[11] and

Beenakkers[12]. During the initial stages of flight, before release of the hormone, energy is derived from trehalose in the haemolymph[14]. A few minutes after the commencement of flight, however, the hormone is released from the intrinsic secretory cells located in the glandular lobes of the corpora cardiaca[13] into the haemolymph; this hormone has two separate sites of action. It causes the release of specific diglycerides from the fat body[15,16] into the haemolymph. These diglycerides, which are transported as lipoproteins in the haemolymph[17], are used by the flight muscle as a source of energy. The adipokinetic hormone also stimulates the utilisation of these lipids by the flight muscle[18]. Consequently, the hormone is important in the maintenance of prolonged flight activity, by enabling the locust to utilise its lipid stores for energy. Mayer and Candy[11] first indicated that adipokinetic hormone was a peptide as its biological activity was destroyed by proteolytic enzymes.

Isolation

Throughout the purification procedure the hormone was located by measuring its lipid-mobilising activity when injected into adult male *Locusta*. Samples to be injected were dissolved in 20 μl simple insect saline (7.5 g NaCl + 0.375 g KCl l^{-1}). Aliquots (5 μl) of haemolymph were withdrawn immediately before and 1 h after injection of the samples and the changes in haemolymph lipid content measured as described previously[13]. A unit of adipokinetic activity was arbitrarily defined as the quantity of hormone that causes the release of 1 μg lipid per μl haemolymph in 1 h.

The hormone was isolated from the glandular lobes of the corpora cardiaca from both the migratory locust, *Locusta migratoria*, and the desert locust, *Schistocerca gregaria*. The material extracted from 3,000 corpora cardiaca (750 μg)
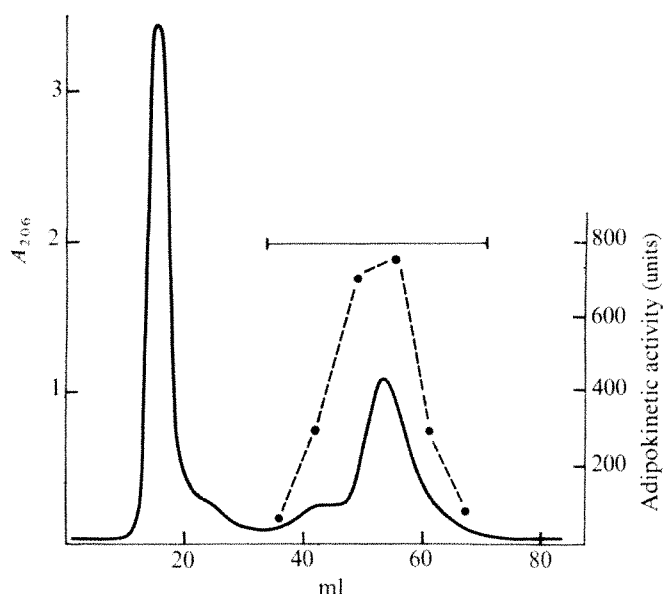
Fig. 1 Separation of an extract of locust corpora cardiaca on G-75-50 glass beads. The residue from a methanol extract of the glandular lobes from 500 locusts was dissolved in 1 ml distilled water and applied to a column (2×25 cm), equilibrated and eluted with water. Flow rate 30 ml h^{-1}. The effluent was collected in 2-ml fractions. The adipokinetic activity of every third fraction, diluted 1:49 with simple saline was measured, taking the mean response from 4 test insects. —, A_{206}; ●, adipokinetic activity; ⊢——⊣, pooled fractions.

has been used to characterise the hormone fully. The glandular lobes were dissected out under simple saline and placed in methanol. The tissue was disrupted using an MSE 100W ultrasonic disintegrator, and the insoluble debris removed by centrifugation. Methanol was evaporated from the supernatant and the residue redissolved in distilled water. This aqueous extract was applied to a controlled-pore glass bead column (G-75-50 beads, Sigma) and the column eluted with distilled water. A profile as shown in Fig. 1 was obtained. The hormone was eluted as a single peak at an elution volume slightly less than the total column volume. The active fractions were pooled and water was removed under reduced pressure at 50 °C. The residue was dissolved in methanol and applied to a silica gel thin-layer plate prewashed with methanol. The chromatogram was developed in isopropanol–water–acetic acid (25:10:1 v/v) which separated out an ultraviolet-absorbing area of R_f 0.67 that contained the adipokinetic activity. The active material was eluted into methanol and subsequently shown to be the pure hormone. The purification procedure, showing the recoveries of active material at each stage, is summarised in Table 1. Overall, approximately half of the original adipokinetic activity in the glandular lobes was recovered by this method.

Between 100 and 250 pmol (calculated from the amino acid content) of the pure hormone were obtained from one *Locusta* compared with 200–250 pmol from one *Schistocerca*. Taking into account the losses during purification it is estimated that the glandular lobes from the corpora cardiaca of one adult *Locusta* contain 0.25–0.55 μg stored hormone, whereas each adult *Schistocerca* contains 0.50–0.80 μg.

Structure elucidation

Analysis of the hormone after acid hydrolysis gave the amino acid composition shown in Table 2a and suggested it was a nonapeptide. The amino acid compositions of the peptides isolated from both *Locusta migratoria* and *Schistocerca gregaria* were found to be almost identical. The hormone did not react with dansyl-chloride[19] and in electro-

phoresis at 50 V cm^{-1} in formate buffer (pH 2.1) on Whatman number 1 paper, it moved with the neutral marker, 1-diaminonaphthalene-sulphonic acid (DNS-OH). It also coelectrophoresed with the neutral marker alanine at pH 6.5 (pyridine–acetate buffer), the peptide being located in each instance by staining with starch–iodide after chlorination[20]. These results suggest that the peptide does not possess any free amino or carboxyl groups.

The blocked N-terminal amino group of the hormone prevented direct determination of the sequence using the dansyl-Edman procedure[19]. Enzymatic digestion of the peptide with thermolysin for 4 h at 37 °C (120 nmol hormone, 1.2 nmol thermolysin in 500 μl 0.067 M NH₄HCO₃ containing 2.5 μmol CaCl₂), caused the appearance of a residue with a free amino group, identified as phenylalanine using the dansyl-chloride procedure. Electrophoresis of the digest at 50 V cm^{-1} on Whatman 3MM paper at pH 6.5 followed by staining, showed two fragments—one acidic with a mobility of +0.55, and one basic with a mobility of −0.29, relative to aspartic acid. Electrophoresis at 50 V cm^{-1} on Whatman number 1 paper at pH 2.1, also indicated the presence of two fragments, one which moved with the neutral marker, DNS-OH, and the other with a mobility of +0.37 relative to dansyl-arginine. Crude molecular weight determinations from these electrophoretic mobilities[21,22] indicated a value of approximately 350 for the acidic (N-terminal) fragment and 700–800 for the basic (C-terminal) fragment. Elution of these peptides from the electrophoretogram gave an 80% yield of the N-terminal fragment and 20% of the C-terminal fragment. The amino acid compositions of the fragments after acid hydrolysis are shown in Table 2b and c.

The N-terminal fragment was degraded sequentially from its C-terminal end by enzymatic digestion, and the digests analysed on the amino acid analyser without previous hydrolysis. Incubation of 9 nmol of this fragment with penicillocarboxypeptidase-SI (ref. 23) for 3.5 h at 37 °C (in 100 μl 0.2 M pyridine-acetate, pH 4.2, containing 0.15 nmol penicillocarboxypeptide-SI) released 8.7 nmol of a ninhydrin-positive substance, identified as asparagine using the dansyl-chloride procedure. Incubation of 8 nmol of the fragment

Table 1 Purification of adipokinetic hormone from glandular lobes of locust corpora cardiaca

Fraction	Total adipokinetic activity (units)	s.e. of estimate (%)	Recovery (%)
Methanol extract	1,158,300	14	100
Pooled column fractions	982,800	5	85
Thin-layer chromatography fraction, R_f 0.67	649,300	12	56

Values are for a preparation from the glandular lobes of 500 locusts.

Table 2 Amino acid composition after acid hydrolysis of adipokinetic hormone and fragments obtained by thermolysin digestion

	a	b	c
		Relative molar quantity in:	
Amino acid	Hormone	N-terminal fragment	C-terminal fragment
Asp	1.82	1.07	1.15
Thr	1.82		1.71
Glu	1.11	1.49	
Pro	0.96		1.00
Gly	1.05		1.71
Leu	1.00	1.00	
Phe	0.97		0.77

Hydrolyses were carried out in 6 M HCl containing 0.1% phenol. Results are expressed relative to Leu = 1.00 for the hormone and N-terminal fragment, and relative to Pro = 1.00 for the C-terminal fragment.
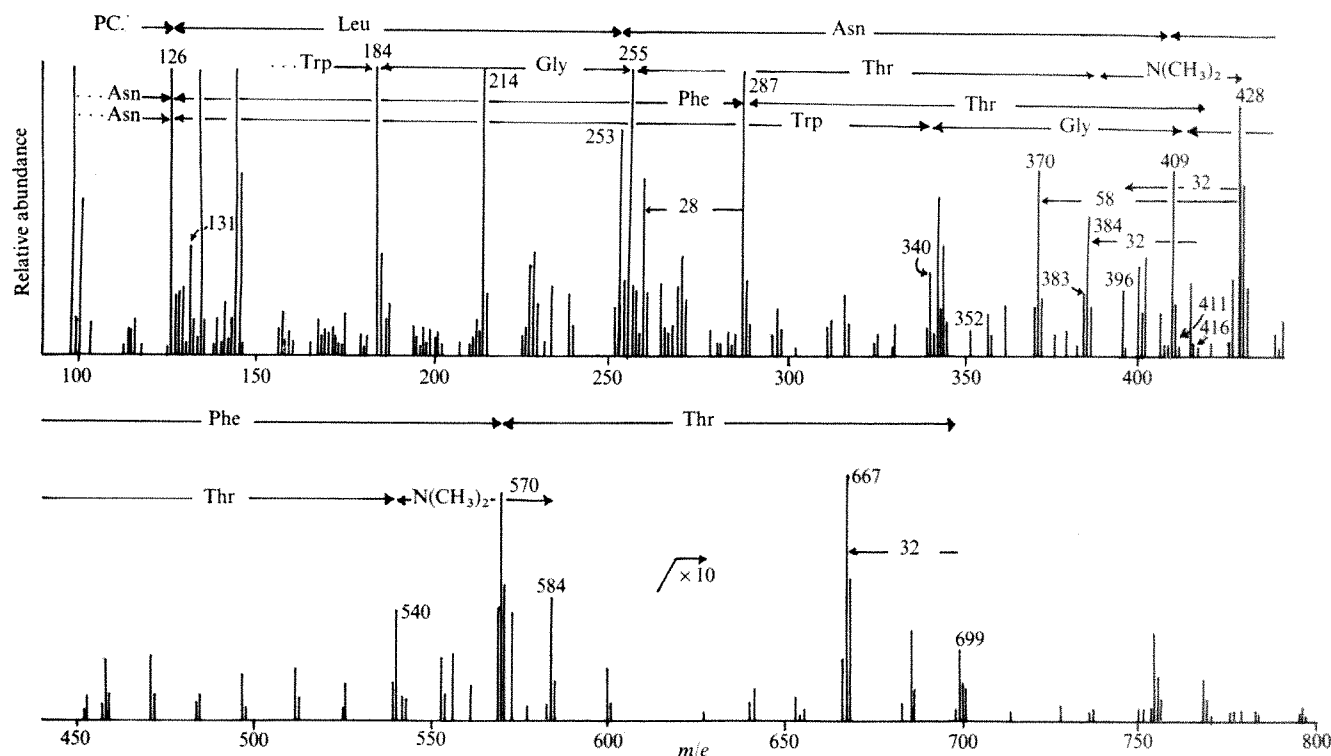
Fig. 2 Mass spectrum at a source temperature of 250 °C of a permethyl derivative of adipokinetic hormone. · · · Asn, · · · Phe and · · · Trp refer to an N–C cleavage fragmentation pathway[25,29].

with (DFP)-treated carboxypeptidase A (prepared by the second procedure of Ambler[24]) for 24 h at 37 °C (in 25 μl 0.2 M NH$_4$HCO$_3$ containing 80 pmol carboxypeptidase A) released 3.8 nmol asparagine and 0.8 nmol leucine. This shows that the fragment has the sequence –Leu–Asn at its C-terminal end, leaving glutamic acid as the N-terminal residue. The N-terminal amino group of this residue is blocked in some way, as it does not react with dansyl-chloride.

The amino acid sequence of the C-terminal fragment was first investigated using the dansyl-Edman degradation procedure. This indicated the sequence Phe–Thr–Pro–Asn– for the first four residues.

So far a partial structure for adipokinetic hormone could be written as: X–Glx–Leu–Asn – – – – Phe–Thr–Pro–Asn–, where X represents some blocking group and the dashed line a probable linkage between fragments. An additional threonine and a glycine residue present in the amino acid analyses of both the intact hormone and the C-terminal thermolytic fragment remained unaccounted for.

Mass spectrometric methods[25] were applied both to the thermolytic fragments and to the intact hormone to determine its overall structure.

Blocked N terminus

Examination of the acetyl-permethyl derivative[26] of a thermolytic digest of adipokinetic hormone was made using the mixture analysis fractional distillation technique[27,28]. Low temperature scans showed an N-terminal sequence fragment at m/e 98 which together with a weak signal at m/e 126 is characteristic of an N-terminal pyroglutamic acid residue.
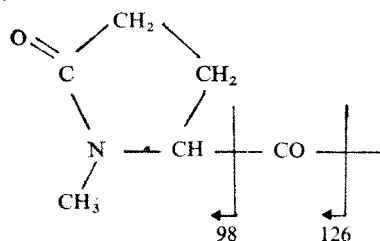
When the associated signal m/e 126 is weak as here, this shows conclusively[25] that the N-terminal residue is a pyrrolidone carboxylic acid (PCA) residue before derivative formation. After these signals, a sequence PCA–Leu was established using a signal at m/e 253.

At higher source temperatures a second component volatilised giving a mass spectrum with signals at m/e 204, N-terminal Phe; at 333, Phe–Thr; at 430, Phe–Thr–Pro; and at 586, Phe–Thr–Pro–Asn.

Other signals in the spectrum provided firm evidence of the presence of an additional amino acid residue, other than those in the amino acid analysis of the acid hydrolysate of adipokinetic hormone (Table 2).

Complete structure and the presence of tryptophan

The mass spectrum described above and all spectra of the intact hormone derivatives (see below) contain strong signals at m/e 144, 184, 214 or 215 (refs 25 and 29), (or the signals of their deuterated equivalents), together with sequence ions, which imply the presence of tryptophan in adipokinetic hormone. Although unexpected, it could readily be placed in the sequence by analysis of the 'thermolytic' and 'intact hormone' spectra. For example, the spectrum described above, interpreted as Phe–Thr–Pro–Asn, contained further signals showing that the asparagine is followed in sequence by tryptophan. In one of the thermolysin digests examined, this tryptophan was C-terminal (m/e 215), although it was not in the studies of the undigested hormone as shown below. The presence of this tryptophan was confirmed and the sequence was extended to the true C terminus by analysis of the spectra of the intact hormone derivatives.

Figure 2 shows the mass spectrum of the acetyl-permethyl derivative[26] of intact adipokinetic hormone, and was interpreted as follows: Abundant sequence ions at m/e 98, 253, 409, 570 and 699 are readily assigned to PCA–Leu–Asn–Phe–Thr as the N-terminal sequence, confirming the blocked N terminus and the presumptive Asn–Phe bond assumed from the thermolysin data. Other major
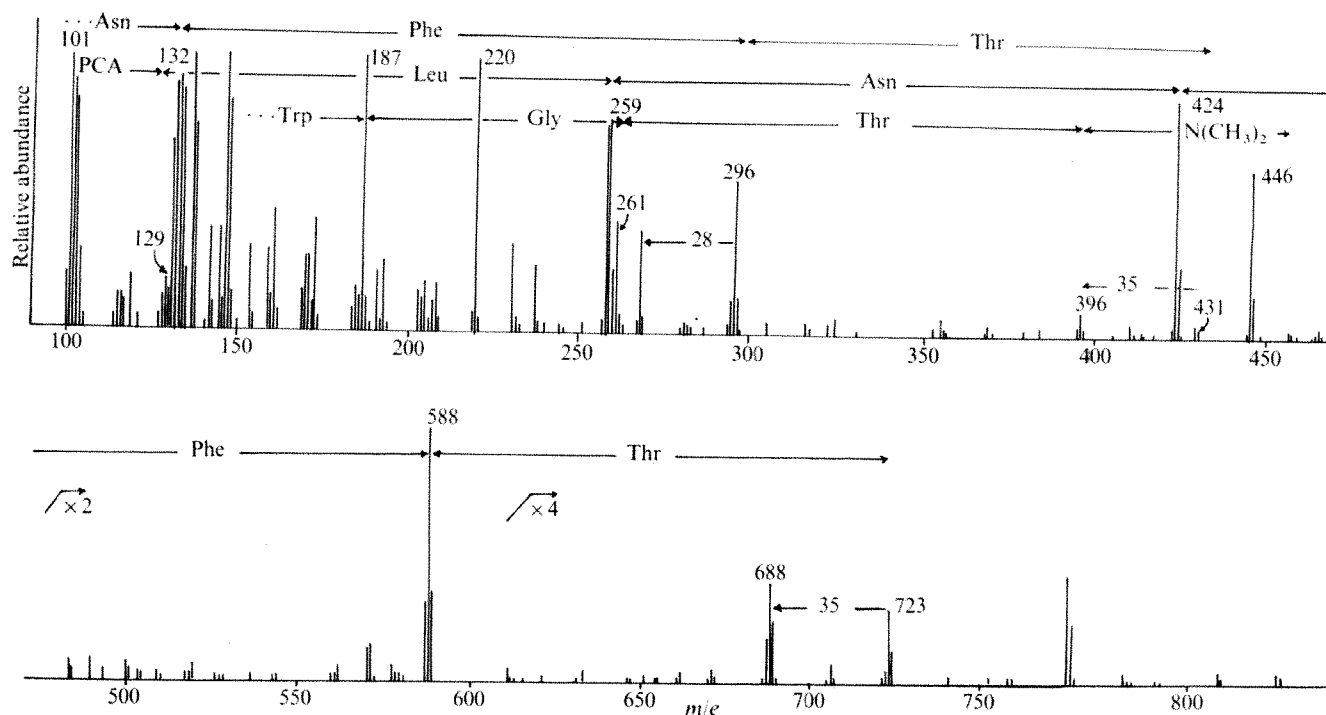
Fig. 3 Mass spectrum at a source temperature of 270 °C of a perdeuteromethyl derivative of adipokinetic hormone. Mass shifts are correlated with signals in Fig. 2 to support the assignments made. For details see text.

signals in the spectrum derive from N–C cleavage pathways[25,27] which originate from signals at m/e 126 (asparagine), 131 (phenylalanine) and 184 (tryptophan).

The most important of these N–C cleavage pathways, as regards interpretation of structure, is that emanating from tryptophan at m/e 184. Testing for amino acid mass differences from m/e 184 the first major unassigned signal is at m/e 255, a glycine mass difference away. That this signal could be due to a threonine residue eliminating its side chain with rearrangement, is unlikely, as neither a threonine sequence ion (m/e 313) nor an associated loss of methanol (m/e 281) are present in the spectrum. Moving from m/e 255 to higher mass, m/e 384 can be assigned to a threonine mass difference, losing the expected 32 mass units (CH₃OH) to give the smaller signal at m/e 352. The reason for m/e 370 not being assigned to a serine residue is that no expected loss of 32 from this signal is observed. After m/e 384 a 44 mass unit difference to m/e 428 is assigned as a C-terminal amide in adipokinetic hormone (an ester would have given a signal at m/e 415). The assignment of m/e 428 to the structure shown below is fully substantiated by the expected losses of methanol (m/e 396) and the threonine side chain (m/e 370).
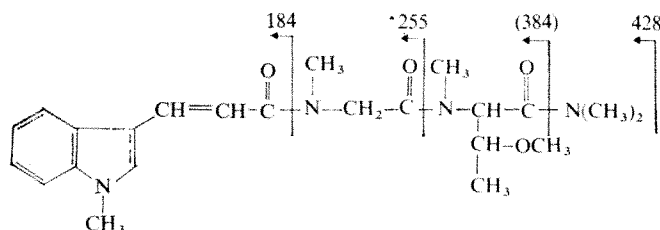
Probably the most powerful confirmation of low resolution structure assignments in mass spectrometry is made by observing the shift in signals associated with the synthesis of deuterated derivatives[30]. The mass spectrum of perdeuteromethyl adipokinetic hormone is shown in Fig. 3. Although the overall appearance of the spectrum is slightly changed by temperature and derivative procedure effects, all the major assignments made in Fig. 2 are vindicated by

the correct mass shift of the corresponding signals. In particular, m/e 187, 261 and 446 in Fig. 3 have moved by 3, 6 and 18 mass units, as would be expected from the structure shown above, and justify the assignment of C-terminal Trp–Gly–Thr–amide. The presence of tryptophan in adipokinetic hormone has since been confirmed both after acid hydrolysis in the presence of 2% thioglycolic acid[31] (30% tryptophan recovery), and after hydrolysis in p-toluene sulphonic acid[32] (50% recovery). Summarising the fragmentation data of Figs 2 and 3 the overall structure of adipokinetic hormone is shown to be: PCA–Leu–Asn–Phe–Thr–Pro–Asn–Trp–Gly–Thr–NH₂. The hormone is therefore a blocked decapeptide and its molecular weight is 1,158.

Biological activity

The biological activity of the fragments produced by thermolysin digestion of the hormone was investigated and compared with the activity of the natural hormone. Figure 4 shows a dose–response curve for pure adipokinetic hormone injected into adult male *Locusta*. At doses above about 5 pmol of hormone the response is near maximal. Doses of the thermolytic digest containing up to 30 pmol of each fragment, however, did not elicit any lipid-mobilising response when injected into locusts. (Thermolysin was inactivated by boiling before injection of the sample.) Consequently, it seems that cleavage of the hormone between its third and fourth residues completely destroys its biological potency, as neither of the resulting fragments possesses any adipokinetic activity.

It is clear that biologically active peptides are components of various control in insects. It was not until relatively recently, however, that any such peptide had been isolated in sufficient quantity for its molecular structure to be determined. The amino acid sequences of a 27-residue peptide from the male fruit fly's accessory gland that influences female mating behaviour[33], and of a pentapeptide that is a proposed neurotransmitter in the cockroach[34], have both been elucidated. To our knowledge the data presented here form the first instance of the complete characterisation of an insect neurohormone. This elucidation of the detailed structure of an insect peptide hormone may be the first step in development of new ways to control
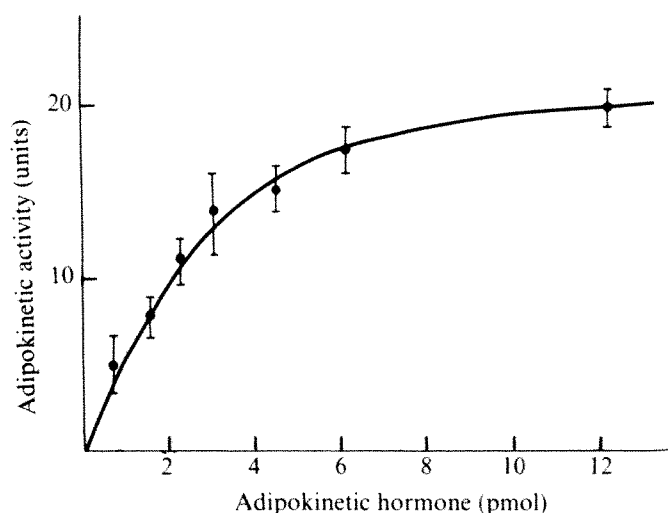
Fig. 4 Dose–response curve for locust adipokinetic hormone. Samples of the pure hormone, dissolved in 20 µl simple saline, were injected into 16-d-old adult male *Locusta*. Each point represents the mean activity ±s.e. from 6 animals.

insect pests. Once the structure–activity relationships are resolved, it may be possible to design compounds which interfere specifically with the normal functioning of the insect neurosecretory system; such interference will d'srupt many morphological, metabolic and physiological events. Indeed it has been shown that treatment with conventional insecticides provokes the release of large quantities of neurohormones, and it is suggested that such release may be the, or a contributory, cause of death[35,36].

The cerebral neurosecretory cell–corpus cardiacum complex in insects is analogous with the hypothalamic–neurohypophysial system in vertebrates[37]. Therefore it is of particular interest that the insect hormone identified here possesses features in common with certain hormones secreted by the hypothalamus. Thus both the N-terminal pyroglutamic acid residue and the C-terminal amide shown for adipokinetic hormone in the present paper, are also seen in thyrotropin-releasing hormone[38,39] and luteinising hormone-releasing hormone[40,41]. A recently characterised crustacean neurohormone, the red-pigment-concentrating hormone from shrimps[42], also possesses these features. Moreover, in spite of its different biological function this shrimp hormone has a similar amino acid sequence to adipokinetic hormone (adipokinetic hormone: PCA–Leu–Asn–Phe–Thr–Pro–Asn–Trp–Gly–Thr–NH_2; shrimp hormone: PCA–Leu–Asn–Phe–Ser–Pro–Gly–Trp–NH_2). Some

early experiments[43,44] in which extracts of insect corpora cordiaca caused a bleaching effect when injected into shrimps emphasises the similarity in structure of these two hormones. Further characterisation of hormones is clearly necessary before it can be ascertained whether close structural relationships exist among all arthropod neurohormones.

The synthesis of adipokinetic hormone is in progress. It is expected that the synthetic peptide will confirm the structure of the hormone postulated here.

1 Goldsworthy, G. J., and Mordue, W., *J. Endocr.*, **60**, 529–558 (1974).
2 Maddrell, S. H. P., in *Insect Neurobiology* (edit. by Treherne, J. E.), 307–357 (North-Holland/American Elsevier, Amsterdam and New York, 1974).
3 Fraenkel, G., Hsiao, C., and Seligman, M., *Science*, **151**, 91–93 (1966).
4 Mordue, W., *J. Insect Physiol.*, **15**, 273–285 (1969); *Gen. comp. Endocr.*, suppl., **3**, 289–298 (1972).
5 Steele, J. E., *Nature*, **192**, 680–681 (1961).
6 Friedman, S., *J. Insect Physiol.*, **13**, 397–405 (1967).
7 Goldsworthy, G. J., *J. Insect Physiol.*, **15**, 2131–2140 (1969).
8 Davey, K. G., *Adv. Insect Physiol.*, **2**, 219–245 (1964).
9 Mordue, W., and Goldsworthy, G. J., *Gen. comp. Endocr.*, **12**, 360–369 (1969).
10 Gersch, M., and Stürzebecher, J., in *Insect Endocrines, Proc. int. Symp. Insect Endocr., Brno* (edit. by Novák, V. J. A., and Sláma, K.), 71–80 (Academia, Praha, 1966).
11 Mayer, R. J., and Candy, D. J., *J. Insect Physiol.*, **15**, 611–620 (1969).
12 Beenakkers, A. M. T., *Gen. comp. Endocr.*, **13**, 492 (1969).
13 Goldsworthy, G. J., Mordue, W., and Guthkelch, J., *Gen. comp. Endocr.*, **18**, 545–551 (1972).
14 Weis-Fogh, T., *Phil. Trans. R. Soc.*, **B237**, 1–36 (1952).
15 Spencer, I. M., and Candy, D. J., *Biochem. Soc. Trans.*, **2**, 1093–1098 (1974).
16 Jutsum, A. R., and Goldsworthy, G. J., *J. Insect Physiol.*, **22**, 243–249 (1976).
17 Mayer, R. J., and Candy, D. J., *Nature*, **215**, 987 (1967).
18 Robinson, N. L., and Goldsworthy, G. J., *J. comp. Physiol.*, **89**, 369–377 (1974).
19 Hartley, B. S., *Biochem. J.*, **119**, 805–822 (1970).
20 Rydon, H. N., and Smith, P. W. G., *Nature*, **169**, 922–923 (1952).
21 Offord, R. E., *Nature*, **211**, 591–593 (1966).
22 Bailey, C. J., and Ramshaw, J. A. M., *Biochem. J.*, **135**, 889–891 (1973).
23 Jones, S. R., and Hofmann, T., *Can. J. Biochem.*, **50**, 1297–1310 (1972).
24 Ambler, R. P., *Meth. Enzym.*, **11**, 155–166 (1967).
25 Morris, H. R., Williams, D. H., Midwinter, G. G., and Hartley, B. S., *Biochem. J.*, **141**, 701–713 (1974).
26 Morris, H. R., *FEBS Lett.*, **22**, 257–260 (1972).
27 Geddes, A. J., *et al.*, *Biochem. J.*, **114**, 695–702 (1969).
28 Morris, H. R., Williams, D. H., and Ambler, R. P., *Biochem. J.*, **125**, 189–201 (1971).
29 Dell, A., and Morris, H. R., *Biochem. biophys. Res. Commun.*, **61**, 1125–1132 (1974).
30 Morris, H. R., Dell, A., Petersen, T. E., Sottrup-Jensen, L., and Magnusson, S., *Biochem. J.*, **153**, 663–679 (1976).
31 Matsubara, H., and Sasaki, R. M., *Biochem. biophys. Res. Commun.*, **35**, 175–181 (1969).
32 Liu, T. Y., *Meth. Enzym.*, **25**, 44–55 (1972).
33 Baumann, H., Wilson, K. J., Chen, P. S., and Humbel, R. E., *Eur. J. Biochem.*, **52**, 521–529 (1975).
34 Starratt, A. N., and Brown, B. E., *Life Sci.*, **17**, 1253–1256 (1975).
35 Maddrell, S. H. P., and Casida, J. E., *Nature*, **231**, 55–56 (1971).
36 Samaranayaka, M., *Gen. comp. Endocr.*, **24**, 424–436 (1974).
37 Hanström, B., *Nature*, **171**, 72–73 (1953).
38 Burgus, R., *et al.*, *Nature*, **226**, 321–325 (1970).
39 Bowers, C. Y., Schally, A. V., Enzmann, F., Bøler, J., and Folkers, K., *Endocrinology*, **86**, 1143–1153 (1970).
40 Matsuo, H., Baba, Y., Nair, R. M. G., Arimura, A., and Schally, A. V., *Biochem. biophys. Res. Commun.*, **43**, 1334–1339 (1971).
41 Burgus, R., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **69**, 278–282 (1972).
42 Fernlund, P., *Biochim. biophys. Acta*, **371**, 304–311 (1974).
43 Hanström, B., *Hormones in Invertebrates* (Clarendon, Oxford, 1939).
44 Thomsen, M., *Biol. Meddr.*, **19**, 1–38 (1943).

3' Non-coding region sequences in eukaryotic messenger RNA

N. J. Proudfoot & G. G. Brownlee

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The sequence A–A–U–A–A–A is present in six different purified messenger RNA molecules (specifically the α- and β-globulin mRNAs of rabbit and human, the immunoglobulin light chain mRNA of mouse (MOPC 21) and the ovalbumin mRNA of chicken) about 20 residues away from the 3'-terminal poly(A) sequence. In addition, a large selection of the 3' non-coding regions of rabbit and human globin mRNAs (both the α- and β- globin mRNAs) are 85% homologous, demonstrating that this region is significantly conserved in evolution.

EUKARYOTIC mRNAs have been shown to contain several surprising structural features. For instance, they contain modified terminal sequences, specifically the 3' terminal poly(A)[1] and the 5' terminal m⁷Gppp or cap structure[2,3], both of which are added to mRNA post-transcriptionally. In addition, mRNA from higher organisms has been shown to contain substantially more sequence than is required to code for protein. A generalised diagram of the organisation of the various regions in mRNA is shown in Fig. 1.

The cap structure and at least a part of the 5' non-coding sequence are thought to be involved in the binding of the mRNA to the ribosome in the correct initiation of protein
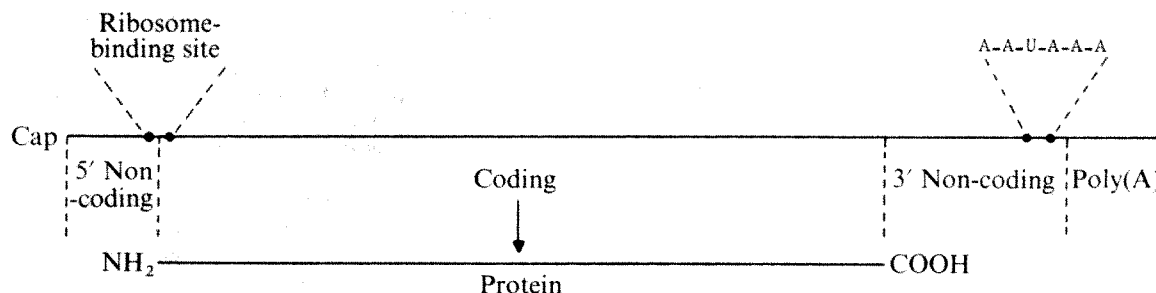
Fig. 1 Generalised diagram of a mRNA molecule. There are, however, exceptions to this general arrangement. For example, histone mRNAs lack poly(A) and polio mRNA lacks Cap[4]. See text for explanation of ribosome binding site and A-A-U-A-A-A.

synthesis[4]. The details of these processes are, however, still unknown. Thus a region of about ten nucleotides, referred to as the 'ribosome-binding site' (Fig. 1) seems to base-pair with a sequence at the 3' end of the 16S ribosomal RNA in *Escherichia coli*[5,6]. Although such a process has yet to be demonstrated in higher organisms, it would seem likely that such an important mechanism would be conserved in evolution.

In contrast to the 5' non-coding region, the function of at least a part of which the sequence is known, there is no direct information on the function of the 3' non-coding region. This is surprising as the 3' non-coding region is longer than the 5' non-coding region—at least with the globin mRNAs, for which reasonably accurate estimates have been made[7]. To find out more about this part of the structure of mRNA we have developed a method that has enabled us to determine the sequence of that part of the 3' non-coding sequence which lies immediately adjacent to the poly(A)[8,9]. By comparing this sequence from a number of different mRNAs we have attempted to locate common features which might indicate regions with particularly important functions. We therefore present here sequences of varying lengths adjacent to the poly(A) of rabbit α and β globins[7,9], human α and β globins (N.J.P. and J. I. Langley, unpublished), mouse immunoglobulin light chain[10] and chicken ovalbumin[11] mRNAs. A comparison of particular interest was that of the sequences of the globin mRNAs in rabbit and man. If these 3' non-coding sequences showed reason-

able evolutionary stability this would argue for their importance functionally.

Sequence homologies

Figure 2 shows the sequences adjacent to poly(A) for the six mRNAs, specifically rabbit α and β globins, human α and β globins, mouse immunoglobulin light chain (MOPC 21) and chicken ovalbumin mRNA. All six mRNAs possess the homologous sequences A-A-U-A-A-A (enclosed in boxes) between 14 and 20 nucleotides from the poly(A) sequence. In addition, human β-globin mRNA and chicken ovalbumin mRNA have a more extensive region of sequence (Fig. 3). In detail a region of 13 nucleotides is completely homologous, save one nucleotide, for the two mRNAs. It is also interesting, although possibly less significant, that the other four mRNA sequences have an additional G residue conserved on the 3' side of the homologous sequence: A-A-U-A-A-A-G (Fig. 2).

Another region of partial sequence homology between the six mRNAs is the sequence directly adjacent to the poly(A). Thus five of the six mRNAs possess the sequence G-C-poly(A). The immunoglobulin light chain mRNA, however, does not possess this sequence. This result is a correction to a previous publication[10] in which a single G residue was incorrectly inserted. Thus mRNA can have different sequences adjacent to poly(A), a result in agreement with other studies[16,17].

30 20 10
a U-G-G-U-C-U-U-U-G-[A-A-U-A-A-A]-G-U-C-U-G-A-G-U-G-A-G-U-G-G-C-poly(A)

30 20 10
b U-G-G-C-U-[A-A-U-A-A-A]-G-G-A-A-A-U-U-U-A-U-U-U-C-A-U-U-G-C-poly(A)

30 20 10
c U-G-G-U-C-U-U-U-G-[A-A-U-A-A-A]-G-U-C-U-G-A-G-U-G-G-G-C-G-G-C-poly(A)

30 20 10
d U-G-C-C-U-[A-A-U-A-A-A]-A-A-A-C-A-U-U-U-A-U-U-U-C-A-U-U-G-C-poly(A)

30 20 10
e A-A-U-A-A-U-U-C-[A-A-U-A-A-A]-G-U-G-A-G-U-C-U-U-U-G-C-A-C-U-U-G-poly(A)

30 20 10
f C-C-U-U-U-A-A-A-U-C-A-U-[A-A-U-A-A-A]-A-A-C-A-U-G-U-U-U-A-A-G-C-poly(A)
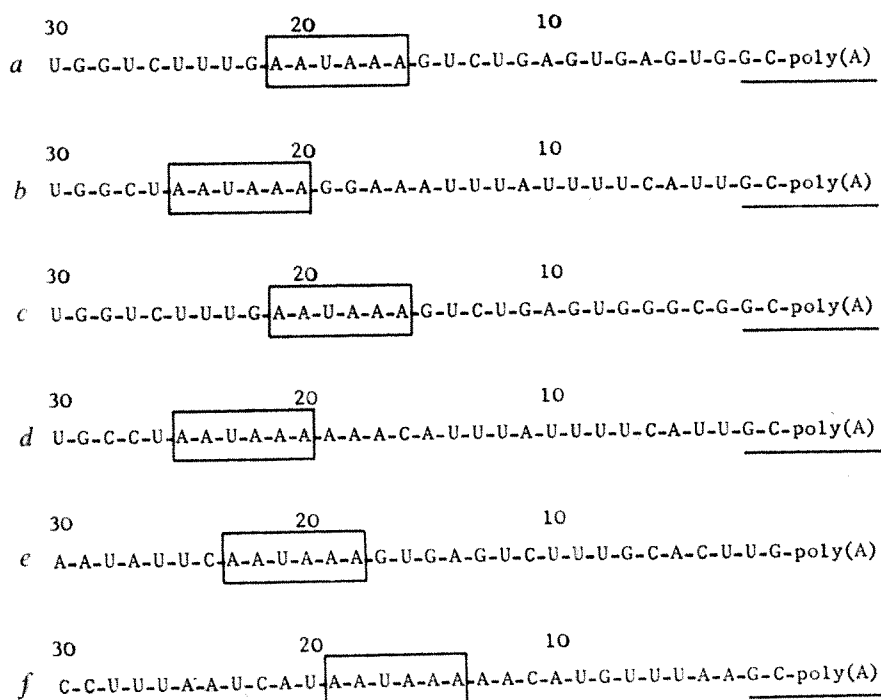
Fig. 2 Comparison of sequences adjacent to poly(A) in six eukaryotic mRNAs: *a*, rabbit α-globin; *b*, rabbit β-globin; *c*, human α-globin; *d*, human β-globin; *e*, mouse immunoglobulin light chain; *f*, chicken ovalbumin mRNAs. The full sequences of the globin mRNAs are shown in Fig. 4, whereas those of the immunoglobulin light chain and chicken ovalbumin mRNAs are presented elsewhere[10,11]. Boxed sequences denote homology, whereas underlined sequences denote partial homology. Numbers indicate distance from poly(A). The sequence analysis of the six mRNAs is described in detail elsewhere (refs 7, 10 and 11, and N.J.P. and J. I. Langley, unpublished). Each mRNA was copied into ^{32}P-labelled cDNA using DNA polymerase I (*E. coli*) in the presence of Mn[2+] and α-^{32}P-labelled deoxyribonucleoside triphosphates with oligo(dT)[10] as primer[8]. In such conditions, short transcripts of the mRNAs were obtained (50–100 nucleotides long) complementary to the 3'-terminal regions of the mRNAs[7]. These cDNA molecules were then digested with endonuclease IV (refs 12 and 13) and the products obtained were sequenced using standard DNA sequencing procedures[13-15]. Note that the mouse immunoglobulin light chain and the rabbit α- and β-globin sequences contain corrections to previously reported sequences[9,10]. These mistakes were detected by improvements in the sequencing methodology[7,11].

Fig. 3 Comparison of a small section of human β globin (*a*) and chicken ovalbumin mRNAs (*b*). Outside box encloses homologous sequence, inside box encloses sequence common to all six mRNAs. Numbers denote distance of that nucleotide from poly(A) sequence.

Comparison of human and rabbit sequences

One direct test as to the functional importance of a particular nucleotide sequence is to establish how conserved or otherwise that sequence is among different species. If the sequence has completely diverged as in the case of satellite DNA[18] its function may be regarded as not sequence specific. Alternatively if the sequence has been very highly conserved, its function can be considered to be highly specific, and central to the overall biological role of the molecule. Such may be the case with the 3′ non-coding regions of globin mRNA. Figure 3 compares the nucleotide sequences determined for our own and other studies for rabbit α- and β-globin and human α- and β-globin mRNAs. As indicated, the sequences of these 3′ non-coding regions are not yet completely known as there is a gap of about 50 nucleotides in both α-globin mRNA sequences and another of 50 or more nucleotides in the β-globin mRNA sequence[7]. Nevertheless, sufficient of the sequence is known to make a preliminary comparison of the data informative.

Both the α- and β-globin mRNA sequences are extensively conserved between human and rabbit. Indeed, of

129 nucleotides from the 3′ non-coding sequences, only 11 substitutions occur. In addition, there are four examples of deletions and insertions of between one and three nucleotides. This may be regarded as 15.5% nucleotide sequence variation between the two species.

One way of assessing the significance of this figure of about 15% is to compare it with the extent of divergence of the respective coding regions. There is, however, insufficient sequence data available at this time to make such a direct comparison possible. Alternatively the coding regions of these mRNAs may be considered in two parts: those nucleotides that on mutation would cause an amino acid replacement, and those that would not (neutral mutations). From amino acid sequence data on rabbit and human globins[21] the number of mutations causing amino acid substitutions between the two species may be calculated as 5.9%. Furthermore, Paddock *et al.*[22] have compared their limited sequence data on rabbit globin mRNA with the sequences determined for human globin mRNA by Marotta *et al.*[23]; and have suggested that the degree of neutral mutation is much larger, as much as 55%. Further sequence data are required to check this figure. From this preliminary conclusion it follows that the 15% divergence shown here in the 3′ non-coding regions is a larger divergence than is allowed in those nucleotides which would change the amino acid sequence, but a smaller divergence than occurs in the degenerate (neutral) nucleotides. The significance of this in terms of function is probably that the 3′ non-coding region is under less selective pressure than the nucleotides specifying amino acid sequence. This is perhaps not surprising. But equally the 3′ non-coding region is much more conserved than are the nucleotides which are degenerate in the triplet code.

The above conclusions are preliminary because the complete 3′ non-coding sequence is unknown. Indeed it is quite

Fig. 4 Comparison of human and rabbit globin mRNA sequences: *a*, α mRNA; *b*, β mRNA. H, Human; R, rabbit; X, base change between human and rabbit; +, addition of sequence; −, deletion of sequence. With α-globin mRNA sequences, termination regions for the two species are also compared and the position of the carboxy-terminal amino acid sequence of α globin (the same for both species, see ref. 23) is positioned. The mRNA sequence connecting the termination regions with the 3′-terminal regions is indicated by a line. The nucleotide sequence determination of the rabbit and human globin mRNA 3′-terminal regions is described in detail elsewhere (ref. 7 and N.J.P. and J. I. Langley, unpublished). The termination sequence for the rabbit α-globin mRNA has been determined by N.J.P. (unpublished), whereas the human α-globin termination sequence has been established by others[19,21].

possible that only the termination region and the sequence adjacent to poly(A) is conserved at the 15% level, whereas there is a greater sequence variation in the rest of the 3' non-coding region.

Discussion and conclusions

The A–A–U–A–A–A sequence present in all six mRNA 3' terminal sequences seems to display the most significant sequence homology between mRNAs. As this sequence has been demonstrated in four different species—human, mouse, rabbit and chicken—it seems likely that it is present in all eukaryotic mRNA, although further work will be necessary to be completely certain. The biological functions of this sequence remain unknown, although one may speculate from its proximity to the 3' end of the mRNA that it is a signal for an enzyme involved either in the synthesis or processing of mRNA from its presumed precursor[4]. The conservation of sequence between the 3' non-coding sequences of human and rabbit globin mRNA further suggests that much of this non-coding sequnce, in addition to the A–A–U–A–A–A sequence, might possess defined sequence-specific functions. As much of the non-coding sequences described above are not homologous for different mRNAs, however, these functions may be specific to individual mRNAs rather than common to all.

The direct sequence analysis of complementary DNA (cDNA) has provided, and no doubt will continue to provide, a powerful method for the sequence determination of eukaryotic mRNA. This method, as with others[23,24], has, however, certain limitations. In particular the 5'-terminal regions of mRNA are particularly inaccessible when priming the synthesis of cDNA with oligo(dT) hybridised to the 3' terminal poly(A). It may be feasible, however, to over-

come this problem by using primers complementary to other regions of mRNA. Also, it is now possible to sequence cDNA that has been inserted into a bacterial plasmid[25–28]. Such DNA sequences are in a double-stranded state and are therefore susceptible to restriction enzyme digestion and the new rapid DNA-sequencing procedure described by Sanger and Coulson[29].

We thank E. M. Cartwright for assistance. N.J.P. is a junior Beit Memorial Fellow.

1. Brawerman, G., *A. Rev. Biochem.*, **42**, 621–642 (1974).
2. Furuichi, Y., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **72**, 1904–1908 (1975).
3. Adams, J. M., and Cory, S., *Nature*, **255**, 28–33 (1975).
4. Perry, R. P., *A. Rev. Biochem.*, (in the press).
5. Shine, J., and Dalgarno, L., *Nature*, **254**, 34–38 (1975).
6. Steitz, J. A., and Jakes, K., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 4734–4738 (1975).
7. Proudfoot, N. J., *J. molec. Biol.* (in the press).
8. Proudfoot, N. J., and Brownlee, G. G., *FEBS Lett.*, **38**, 179–183 (1974).
9. Proudfoot, N. J., and Brownlee, G. G., *Nature*, **252**, 359–362 (1974).
10. Milstein, C., Brownlee, G. G., Cartwright, E. M., Jarvis, J. M., and Proudfoot, N. J., *Nature*, **252**, 354–359 (1974).
11. Cheng, C. C., *et al.*, *J. molec. Biol.* (in the press).
12. Sadowski, P. D., and Bakyta, I., *J. biol. Chem.*, **247**, 405–412 (1972).
13. Galibert, F., Sedat, J. W., and Ziff, E. B., *J. molec. Biol.*, **87**, 377–407 (1974).
14. Sanger, F., Donelson, J. E., Coulson, A. R., Kössel, H., and Fischer, D., *J. molec. Biol.*, **90**, 315–333 (1974).
15. Ling, V., *J. molec. Biol.*, **64**, 87–102 (1972).
16. Nichols, J. L., and Eiden, J. J., *Biochemistry*, **13**, 4629–4633 (1974).
17. Winters, M. A., and Edmonds, M., *J. biol. Chem.*, **248**, 4763–4768 (1973).
18. Southern, E. M., *J. molec. Biol.*, **94**, 51–69 (1975).
19. Seid-Akhaven, M., Winter, W. P., Abramson, P. K., and Rucknagel, D. L., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 882–886 (1976).
20. Clegg, J. B., Weatherall, D. J., and Milner, P. F., *Nature*, **234**, 337–340 (1971).
21. Dayhoff, M. O., in *Atlas of Protein Sequence and Structure*, **4**, D-42 and D-54 (National Biomedical Research Foundation, Silver Spring, 1969).
22. Paddock, G. V., *et al.*, *Proc. natn. Acad. Sci. U.S.A.* (in the press).
23. Marotta, C. A., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **71**, 2300–2304 (1974).
24. Poon, R., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **71**, 3502–3506 (1974).
25. Rougeon, F., Kourilsky, P., and Mach, B., *Nucleic Acids Res.*, **2**, 2365–2378 (1975).
26. Rabbitts, T. H., *Nature*, **260**, 221–225 (1976).
27. Efstratiadis, A., Kafatos, F. C., Maxam, A., and Maniatis, T., *Cell*, **7**, 279–288 (1976).
28. Salser, W., *et al.*, in *Progress in Nucleic Acid Research and Molecular Biology*, **17**, (edit. by Cohn, W. E.), (Academic, New York, in the press).
29. Sanger, F., and Coulson, A. R., *J. molec. Biol.*, **94**, 441–448 (1975).

letters to nature

Effects of physical adsorption on porous interstellar grains

SOLID grains in space are likely to be highly porous in character. Data from meteorites[1], micro-meteoritic trajectories[2], as well as from comets[3] provide indirect evidence that Solar-System grains are porous and a similar porosity is to be expected for interstellar grains. This would have important consequences for interstellar molecular hydrogen formation, and for the expected densities of noble gases in meteorites:

The effects of porosity in laboratory materials such as zeolites, silica gels and active carbons have been studied for many years[4], mainly with a view to exploiting their properties as adsorbents and molecular sieves. Adsorbant pores are voids in solids which usually communicate with each other. Pore sizes range from 1,000 Å (macropores) to a few Å (micropores). The distribution of pore sizes to be expected in astrophysical solids depends on the mechanism by which they form. The sticking together of separate refractory grains embedded in a matrix of volatile ices or polymers which subsequently partially evaporate because of heating, could result in macroporous structures, as shown in micro meteorites (mass densities ~0.3 g cm^{-3}). Microporosity with pore dimensions < 10 Å may result from the thermal dehydration of originally hydrated mineral crystals (for example zeolites or silica gels).

For macroporous grains the total effective adsorption area could be increased by an order of magnitude, but the physisorp-

tion energy remains unaltered. Microporosity, on the other hand, leads to enhancement of effective physisorption energy by a factor of ~5 in a typical case (Fig. 1, ref. 4), together with an increase of the adsorption area of a grain by a few orders of magnitude. The surface area per unit mass for non-porous spherical grains of radius a, density s is $3/as$. For typical values $a = 1.5 \times 10^{-5}$ cm, $s = 2$ g cm^{-3}, this gives a 'specific adsorption area' 10^5 cm^2 g^{-1}. In the case of a microporous material (for example zeolites) an average value for the specific adsorption area is 1.5×10^7 cm^2 g^{-1}, a factor 150 greater than for non-porous grains. This enhancement is interpreted as resulting from a volume-filling effect. The consequences of these properties for the formation of hydrogen molecules in interstellar space and for the retentive siting of noble gases in meteorites are examined below.

The condition for efficient H_2 formation on a grain surface is that the residence time of an adsorbed atom exceeds the mean time interval between successive arrivals of H atoms. This gives a maximum grain temperature

$$T_m = \frac{q}{\ln(\alpha \langle v \rangle n_H \sigma \tau_0)}$$

where q is the binding energy (Kelvin) of an H atom, α is the sticking coefficient, $\langle v \rangle$ is the mean H atom velocity, σ is the grain cross-sectional area, and τ_0 is a characteristic desorption lifetime. Early estimates[5] including the effect of zero-point

energy in the adsorption energy, with $q = 200$ K, yielded $T_m \simeq 6$ K for typical interstellar conditions. This value of grain temperature is unacceptably low. Even the later calculations of Hollenbach and Salpeter[6] which give $q \sim 400$ K (for hydrogen adsorbed on water ice) and $T_m \sim 12$ K allow molecular formation to occur only in clouds where dust shielding is significant. To increase T_m they proposed sites of greater binding on the grain, because of the presence of impurities and dislocations, with $q \sim 2,000$ K. For the limited number of these sites they chose, T_m increased to 40 K. Microporosity can, however, lead to an increase in q at all sites in all micropores. An increase by a factor of ~ 5 gives maximum grain temperatures of ~ 30 K. Even higher values of T_m are likely since the value $\tau_0 = 6 \times 10^{-13}$ s used in earlier calculations is based on the Frenkel theory of adsorption for non-porous surfaces. An increase of τ_0 by several orders of magnitude, likely for porous materials, would lead to $T_m \gg 30$ K. No dust shielding is then required, and H_2 recombination could take place on relatively 'hot' grains, even in intercloud regions. Such an effect could be important in accounting for recent Copernicus observations[7].

Porous grains will have a major role in the origin of primordial noble gases in meteorites. An outstanding problem concerning the origin of noble gases in meteorites is the enormously large nebular pressures needed to account for the observed abundances in carbonaceous chondrites, if one assumes equilibrium solubility models[8,9]. Using the experimental data of solubilities in magnetite, the pressures needed to obtain the observed amounts of rare gases are 5–8 orders of magnitude larger[10] than might be reasonably estimated[11]. To overcome this difficulty it has been hypothesised[8] that the equilibrium solubility of noble gases are 5–8 orders of magnitude larger in the minerals where these gases are found than in magnetite. Alternatively, it has been proposed that the noble gases may be trapped by physical adsorption, the need for a hypothetical increase in solubility thus being avoided. The nebular pressures required for such adsorption are marginally plausible[12] within the framework of detailed numerical models for the condensation of the solar nebula[11]. Furthermore, the observed elemental pattern of the noble gases is well reproduced by an adsorption process.

This model, however, based on measurements made on the Allende meteorite, suffers two major drawbacks[10]. Extensive heating[13] and etching[14] experiments on several meteorites indicate that the associated gases are firmly bound, and seem to be uniformly distributed throughout the host crystal. These facts may be taken to imply that the role of adsorption in the conventional sense is minimal.

The presence of micropores will alter this picture. The retentivity will increase enormously from an increase in the physisorption energy. Further, the micropores will allow the gases to fill the entire volume of grains as opposed to some specific surfaces alone. Of course, the evidence gathered so far from meteorites indicates merely the existence of pores. The spectrum of the radii of the pores has not yet been determined. The above physical arguments strongly suggest, however, that the minerals containing the noble gases will be highly microporous. An experimental determination of the degree of microporosity of chondrites is thus timely and highly desirable.

We thank the SRC for financial support.

H. ABADI
PUSHPA JOSHI
S. RAMADURAI
N. C. WICKRAMASINGHE

Department of Applied Mathematics and Astronomy,
University College,
Cardiff, UK

1 Stacey, F. D., Loveking, J. F., and Parry, L. G., *J. geophys. Res.*, **66**, 1523 (1961).
2 Millman, P. M., in *From Plasma to Planet, Nobel Symposium*, **21** (edit. by Elvius, A.), 157 (Almquist and Wiksell, Stockholm, 1972).
3 Vanysek, V., and Wickramasinghe, N. C., *Astrophys. Space Sci.*, **33**, L.19 (1975).

4 Dubinin, M. M., *Prog. Surf. Membrane Sci.*, **9**, 1 (1975).
5 Knapp, H. F. P., Van der Meijenberg, C. N. J., Beenaker, J. J. M., and Van de Hulst, H. C., *Bull. astr. Insts Neth.*, **18**, 256 (1966).
6 Hollenbach, D., and Salpeter, E. E., *Astrophys. J.*, **163**, 155 (1971).
7 York, D. G., *Astrophys. J.*, **204**, 750 (1976).
8 Lewis, R. S., Srinivasan, B., and Anders, E., *Science*, **190**, 1251 (1975).
9 Grossman, L., and Larimer, J. W., *Rev. Geophys. Space Phys.*, **12**, 71 (1974).
10 Lancet, M. S., and Anders, E., *Geochim. cosmochim. Acta*, **37**, 1371 (1973).
11 Cameron, A. G. W., and Pine, M. R., *Icarus*, **18**, 277 (1973).
12 Fanale, F. P., and Cannon, W. A., *Geochim. cosmochim. Acta*, **36**, 3919 (1972).
13 Black, D. C., *Geochim. cosmochim. Acta*, **36**, 377 (1962).
14 Jeffery, P. M., and Anders, E., *Geochim. cosmochim. Acta*, **34**, 1175 (1970).

Photon correlation study of stellar scintillation

DURING the past decade photomultiplier-tube technology and signal processing using integrated circuits have progressed to the point at which the analysis of fluctuating light signals can now be carried out digitally, in real time, with accuracy close to the theoretical limit. These 'photon correlation' techniques have so far found their main application in laser light-scattering spectroscopy[1] and laser doppler velocimetry[2]. The techniques need not, however, be limited to applications involving lasers and can be used for the analysis of fluctuating optical signals of any origin (provided the fluctuation time is $\geq 10^{-8}$ s). Here we report a preliminary study of the statistical and temporal correlation properties of stellar scintillation, taking advantage of the fast response time and high sensitivity of equipment now standard for laser light-scattering applications.

The twinkling of stars is, of course, a phenomenon of long standing interest and a continuing source of controversy. It is caused by the presence of fluctuations in the refractive index of the Earth's atmosphere, but the exact nature and role of these fluctuations seems to be in some dispute[3,4]. The twinkling phenomenon is associated with a fluctuation in amplitude, or scintillation, of the optical intensity incident on the eye. This is a potential source of information and can be used as a remote probe of the atmosphere just as radio-star scintillation has been used to investigate the ionosphere[5]. There is indeed a good deal of current interest in this aspect of the phenomenon[6,7]. Unfortunately, it is also a form of noise which may limit the resolution of astronomical telescopes. The current development of stellar speckle interferometry[8] as a means of overcoming conditions of bad seeing therefore provides further motivation for a quantitative characterisation and understanding of stellar scintillation.

Observations were made in late January, 1976, at RSRE, Malvern (52°8'N, 2°20'W). The results on Sirius (α Canis Majoris) reported here were obtained between 2030 and 2200

Fig. 1 Normalised intensity correlation function of scintillations in light received from Sirius, combining the results of several measurements taken over a 1.5-h period, not corrected for background. *a*, 5-nm filter; *b*, without filter.
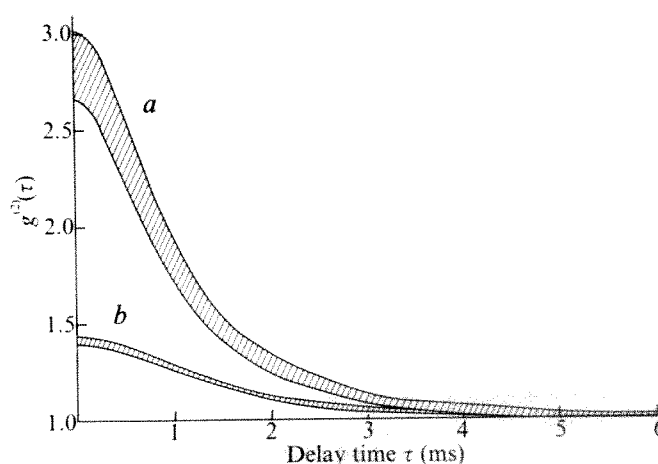
Table 1 Normalised factorial moments $n^{[m]}$ of photon-counting distribution, that is, normalised moments of the intensity fluctuation distribution[1]

m	$n^{[m]}$		
	Uncorrected*	Corrected†	Gaussian‡
2	2.74 ± 0.13	2.89 ± 0.14	2
3	16.9 ± 2.2	18.8 ± 2.5	6
4	191 ± 47	223 ± 55	24
5	$3,226 \pm 1,220$	$3,942 \pm 1,490$	120

*Moments of measured distribution.
†Moments corrected for presence of background light ($\sim 4\%$ total intensity) assumed to be non-fluctuating on time scale of interest.
‡For a Gaussian field, $n^{[m]} = m!$ (ref. 1).

on January 20 with the star coordinates near 165° azimuth and 21° elevation. Qualitatively similar results were obtained with Vega (α Lyr) during the same period. A blustery westerly airstream covered the region at this time with wind speeds of the order of 10–15 knots near the ground. An ITT FW130 (S20) photomultiplier, preceded by an 18.5-cm focal length telephoto lens and a filter centred at 443 nm with bandwidth 5nm and 20% transmission, was mounted about 40 feet above the ground on a radar dish programmed to track the star. The star was imaged with $f/22$ optics (0.84-cm diameter aperture) on to a field stop of diameter 1 mm; this was followed by a dispersing lens which distributed the received light over the photocathode of the detector. The photomultiplier output was thus a measure of the light intensity falling on an area of about 0.55 cm². Photon detection rates were typically greater than 2.5×10^4 s^{-1} when the star was in the field of view and of the order of 10^3 s^{-1} when it was not, made up of 10^2 s^{-1} room temperature dark count and 9×10^2 s^{-1} sky background. The optical assembly, including telephoto lens, viewing system, phototube and amplifier–discriminator circuitry, and the digital correlator used to measure photon statistics and correlations, were standard commercial items (Malvern Instruments, UK). For correlation measurements the correlator was operated in the single-scaled mode[9], which provided a direct measurement of the normalised intensity correlation function, $g^{(2)}(\tau)(= \langle I(0)I(\tau)\rangle/\langle I\rangle^2$ where I is the light intensity and τ the delay time), whatever the statistics of I.

Fig. 2 Probability distribution of photon detections. Experimental points would lie on the straight line if the received light field were Gaussian distributed with the same mean of about 1 count per 40-μs sample time.
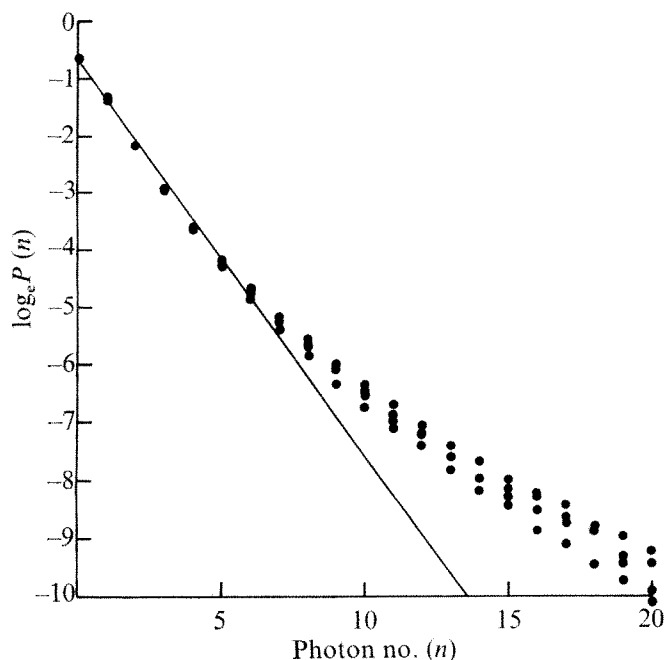
Figure 1a shows a composite correlation function obtained from several measurements of 1–2-min duration with sampling times varying from 20 to 200 μs. A band roughly 2 s.d. wide indicates run-to-run variations which arise from both finite counting times and non-stationary atmospheric conditions. The normalised contrast

$$\lim_{\tau \to 0} g^{(2)}(\tau) - 1$$

corrected for background, is about 1.9, significantly greater than the value of unity expected for Gaussian light. The non-Gaussian nature of the fluctuations is also reflected in the higher-order intensity moments (Table 1) and the photon-counting probability distribution (Fig. 2). A considerably reduced contrast of ~ 0.42 was observed, however, when the colour filter was removed (Fig. 1b). The width at half-height of the correlation function is about 10^{-3} s implying a characteristic fluctuation time much shorter than the commonly accepted value for the response time of the eye. Taking this to be of order 1/25 s, however, a r.m.s. contrast of about 10% after temporal integration of the whole visible spectrum is indicated or 20% if integration is restricted to a 5-nm band. Since the eye perceives fluctuations in colour, the latter figure would seem to be appropriate for comparison and not inconsistent with visual observations, although physiological effects due to image motion on the retina cannot be ruled out[10].

To our knowledge non-Gaussian fluctuations giving contrast in excess of unity have not been observed in any previous measurements of stellar scintillation[11], probably because in earlier experiments averaging of the intensity pattern occurred through the use of large apertures, insufficiently narrow filters or simply slow response time instrumentation. High contrast intensity patterns are known to be produced in certain conditions when an infinite plane wave is scattered by a random phase screen or inhomogeneous medium[12]. In particular, a random-phase screen characterised by a single length scale and introducing path differences greater than about one-third of a wavelength produces large intensity fluctuations in the region of focusing determined by the lens-like behaviour of individual refractive index inhomogeneities[13]. Our results seem to be consistent with recent theoretical predictions based on this interpretation of the origin of stellar scintillation[14]. The phenomenon of focusing, however, is highly model dependent[12,15] and further experiments are essential if the role of the various turbulent layer structures present in the atmosphere is to be fully elucidated.

Finally, it is worth noting that the shadow patterns observed in the receiving plane of large aperture telescopes[16] are, in conditions of high contrast, probably smeared-out caustics. The presence of many of these within the aperture would indicate that light from many (strong) atmospheric scatterers was entering the optical system leading to near-Gaussian speckle in the image[8].

We thank Flt Lt R. C. Perry (RAF) and his staff for assistance in operating the radar dish, and members of the Applied Optics Section at Imperial College and Dr H. E. Butler of the Royal Observatory, Edinburgh, for discussions.

E. Jakeman
E. R. Pike
P. N. Pusey

*Royal Signals and Radar Establishment,
Malvern, Worcestershire WR14 3PS, UK*

1 *Photon Correlation and Light Beating Spectroscopy* (edit. by Cummins, H. Z., and Pike, E. R.), (Plenum, London and New York, 1974).
2 Abbiss, J. B., Chubb, T. W., and Pike, E. R., *Opt. Laser Technol.*, **6**, 249–261 (1974).
3 Young, A. T., *Appl. Optics*, **8**, 869–885 (1969).
4 Vernin, J., and Roddier, F., *J. opt. Soc. Am.*, **63**, 270–273 (1973).
5 Singleton, D. G., *J. atmos. terr. Phys.*, **32**, 187–208 (1970).
6 Rocca, A., Roddier, F., and Vernin, J., *J. opt. Soc. Am.*, **64**, 1000–1004 (1974).
7 Ochs, G. R., Clifford, S. F., Lawrence, R. S., and Wang, T-h, *Natn. oceanic atmos. Admin. Tech. Rep.*, ERL297–WPL30 (1974).
8 Labeyrie, A., *Nouv. Rev. d'Opt.*, **5**, 141 (1975).
9 Jakeman, E., Oliver, C. J., Pike, E. R., and Pusey, P. N., *J. Phys. A.*, **5**, L93–96 (1972).

10 Minnaert, M., *Nature*, 165, 663–664 (1950).
11 See for example Burke, J. J., *J. opt. Soc. Am.*, 60, 1262–1264 (1970).
12 Prokhorov, A. M., Bunkin, F. V., Gochelashvily, K. S., and Shishov, V. I., *Proc. IEEE*, 63, 790–811 (1975).
13 Bramley, E. N., and Young, M., *Proc. IEE, Lond.*, 114, 553–556 (1967).
14 Taylor, L. S., and Infosino, C. J., *J. opt. Soc. Am.*, 65, 78–84 (1975).
15 Jakeman, E., and McWhirter, J. G., *J. Phys. A.*, 9, 785–797 (1976).
16 Mikesell, A. H., Hoag, A. A., and Hall, J. S., *J. opt. Soc. Am.*, 41, 689–695 (1951).

Q in cosmology

THE following example in classical mechanics yields the key equations of relativistic cosmology, including the new equations by Harrison[1]. The argument given here develops that of an earlier note[2].

A set of n gravitating particles of masses m_i are at positions $r_i(0) = a_i$ relative to a fixed origin at the time $t = 0$, when an explosion gives a radial motion $r_i(t) = R(t)a_i$ to all the particles. The scale factor R is the same for all the particles. The kinetic energy of the system at time t is a sum of terms of the type $\frac{1}{2}m_iv_i^2$, and the gravitational potential energy is a sum of terms of the form ($|a_i - a_j| \equiv a_{ij}$)

$$-G\frac{m_i m_j}{r_{ij}(t)} = -G\frac{m_i m_j}{a_{ij}}\frac{1}{R(t)}$$

Accordingly, the energy equation of the system is

$$E = A\dot{R}^2 - B/R, \quad A \equiv \frac{1}{2}\Sigma_i m_i a_i^2, \quad B \equiv \frac{G}{2}\sum_{i \neq j}\frac{m_i m_j}{a_{ij}} \quad (1)$$

To attain a formal analogy with relativistic cosmology, write

$$E/A \equiv -kc^2, \quad B/A \equiv C,$$

and add a cosmological energy term $A\lambda R^2/3$. This gives the following energy equation and its first and second time derivatives

$$\frac{\lambda}{3} = H^2 + \frac{kc^2}{R^2} - \frac{C}{R^3} \qquad (H \equiv \dot{R}/R) \quad (2)$$

$$\frac{\lambda}{3} = \frac{C}{2R^3} - qH^2 \qquad (q \equiv -\ddot{R}R/\dot{R}^2) \quad (3)$$

$$\frac{\lambda}{3} = QH^2 - \frac{C}{R^3} \qquad (Q \equiv R^2\dddot{R}/\dot{R}^3) \quad (4)$$

The importance of the Q parameter has already been noted[1]. One obtains the basic cosmological differential equation in two forms which are independent of the experimentally little known parameter λ, by equating equations (2) and (3) or by equating (2) and (4)

$$\frac{kc^2}{R^2} = \frac{3C}{2R^3} - (q+1)H^2 \quad (5)$$

$$\frac{kc^2}{R^2} = (Q-1)H^2 \quad (6)$$

Adding twice (3) to (4) and taking (3) from (4):

$$\lambda = (Q-2q)H^2, \frac{3C}{2R} = (Q+q)H^2 \quad (7,8)$$

In a Newtonian context, the system is ever expanding if E is positive; that is, if k is negative. By equation (6), the appropriate condition is $Q < 1$, and for $\lambda = 0$ models, it is equivalent to $q < \frac{1}{2}$ (equation (7)). The inequalities are reversed if the system is to collapse again. Note that only if one puts (ρ is the mass density)

$$C = \frac{8\pi}{3}G\rho R^3 \quad (9)$$

do these become the equations given by Harrison[1]. The simplicity of the example given here is illuminating: the role of q, which determines whether a ($\lambda = 0$)-model is closed or open, can clearly be taken over by Q if $\lambda \neq 0$. A continuous distribution of gravitating matter, which is needed to obtain equation (9) in general relativity, and which was used in ref. 1, is clearly not needed here. A discrete distribution of galaxies with sufficient symmetry to ensure the continued validity of Hubble's Law is all that is required. It is to be emphasised that in Newtonian cosmology it is possible to give discussions of interacting particles and discrete matter distributions which give Hubble's Law, though the corresponding general relativistic problems remain unsolved.

This possibility has been noted already[2,3], but no example has been given before. The simplest distribution to verify this possibility is a set of eight galaxies, each of mass m, on the vertices of a cube of side $2a_0$, which are given velocities, v, radially away from the centre of the cube. Because of the seven other galaxies each galaxy is then subject to a deceleration in the direction of the centre. Its magnitude is

$$\frac{D}{2a(t)^2}, \quad D \equiv \left[1 + \frac{\sqrt{2}}{2} + \frac{\sqrt{3}}{9}\right]\frac{Gm}{2}$$

where $2a(t)$ is the side of the cube at time t. If the escape velocity $v_{esc} = \sqrt{(D/a_0)}$ has been exceeded, the Hubble parameter is

$$\frac{\theta^3(1+w^2)^{\frac{1}{2}}}{Dw^3}, \qquad \theta \equiv (v^2 - v_{esc}^2)^{\frac{1}{2}}$$

$$w \equiv \theta\sqrt{[a(t)/D]}.$$

Essentially, a topic for so far undiscovered theorems in Newtonian physics is raised here: the classification of (usually unstable) particle configurations which, for a given interaction, are consistent with a Hubble type expansion; and expressions for the Hubble parameter should also be found in each case.

P. T. LANDSBERG

Department of Mathematics,
The University,
Southampton SO9 5NH, UK

1 Harrison, E. R., *Nature*, 260, 591 (1976).
2 Landsberg, P. T., *Nature*, 242, 104 (1973).
3 Landsberg, P. T., *Nature*, 244, 66 (1973).

Anomalous M_1 tide at Lagos

I HAVE recently described[1] the phenomenal appearance in the seas off western Europe of a tidal component with a period of exactly one mean lunar day, namely the "M_1 tide". Although only a few mm in amplitude, it is about ten times larger than one would expect from comparison with ordinary diurnal tides, and suggests a resonant response of

part of the Atlantic Ocean to forcing functions which are symmetrical with respect to the Equator. The forcing function for the ordinary diurnal tides is antisymmetric, and its components are relatively suppressed, enhancing the appearance of M_1 in this sea area.

Most of the places for which I showed evidence of an enhanced M_1 tide, however, were in the shelf seas surrounding the British Isles. The one notable exception was Terceira (Azores), where the results were somewhat marginal. It might therefore have been suggested that the phenomenon is not really oceanic but is confined to the north-west European shelf, although I gave some reasons why this seemed unlikely. I now describe results which confirm the oceanic scale of this effect.

Resolution of the true M_1 tide requires at least 5 yr, and preferably 9 yr of data. A tidal record of more than 9 yr duration has been obtained from Lagos on the south Portuguese coast (near Cape St Vincent; $37°06'N, 8°40'W$), through the cooperation of the Instituto Geografico e Cadastral at Lisbon. The tide gauge, being maintained for geodetic purposes, is excellent in quality, and analysis of the records shows them to have very low noise level and a virtual absence of nonlinear effects. The true M_1 tide stands out clearly from its spectral background, with an amplitude of 5 mm and Greenwich phase lag G of $245°$, representing an admittance to the normalised generating function[2] of 1.23. By contrast, the major component of the ordinary diurnal tide, at a frequency close to 1 cycle per lunar day but actually $1/8.85$ cycles yr^{-1} greater than that of M_1, has amplitude 4 mm, $G=5°$, representing an admittance to its generating function of only 0.19. Lagos is far enough from Britain not to be affected by resonances in British shelf seas, should they exist, so a significant part of the north-eastern Atlantic must be involved. Also, a definite northwards progression of phases across the eastern oceanic shelf edge now emerges, as seen by the following list of phase lags G for the M_1 tide: Lagos, $245°$; Brest, $268°$; Scilly, $276°$; Malin, $312°$; Stornoway, $308°$; Lerwick, $330°$.

Apart from a slight anomaly at Stornoway, the trend of the phases is in keeping with the progression of the normal mode of the Atlantic/Indian ocean system at 23.5-h period, computed by Platzman[3], which I cited earlier[1] as an explanation of the M_1 phenomenon. Platzman's mode progresses in phase by $45°$ roughly between the positions of Lisbon and the Faeroes. G for Terceira is $275°$, about the same as Scilly, indicating a cotidal line stretching south-west from Scilly. This, together with the reduced amplitude at Terceira, again supports Platzman's modal diagram.

<div style="text-align:right">D. E. Cartwright</div>

Institute of Oceanographic Sciences,
Bidston, Birkenhead,
Merseyside L43 7RA, UK

1 Cartwright, D. E., *Nature*, **257**, 277–280 (1975).
2 Cartwright, D. E., and Tayler, R. J., *Geophys. J. R. astr. Soc.*, **23**, 45–74 (1971).
3 Platzman, G. W., *J. Phys. Oceanogr.*, **5**(2), 201–221 (1975).

Diachronism in Old World alluvial sequences

THE diachronism of an environmental change—that is, the extent to which its age varies with location—is a useful clue to its origin[1]. In the case of alluvial sequences, the time taken for filling and trenching to work their way through a drainage basin is likely to obscure chronological differences between basins unless they are sufficiently distant from each other for any regional trend to become apparent. I report here ^{14}C dates for alluvial units in south-eastern Spain, and suggest that they are consistent with the effects of latitudinally diachronous

Fig. 1 Location of Spanish sampling sites discussed in text.

deposition throughout the Mediterranean Basin and Near East.

The Mundo is a tributary of the Segura[2], which drains into the Mediterranean south of Alicante. Three Quaternary valley deposits can be identified in the reaches shown in Fig. 1. The oldest, for which I propose the name Ayna Tufa, has a maximum thickness of 17 m above Ayna, where it overlies Mesozoic limestone and grades upstream into a well stratified marl. The tufas and marls are surmounted by a deposit which takes the form of cemented scree at Ayna and of a valley fill below Ayna and in many other valleys in the area. This unit, for which the name Elche Formation is proposed, attains its maximum observed thickness of 11 m in the Rambla de Dilar; it is generally characterised by subangular gravel in a red, sandy matrix and includes horizons strongly cemented by calcium carbonate. Channels cut into the Elche sediments have been partially filled, to a maximum observed depth of 2.5 m, by a well bedded deposit composed of silts, sands and gravels for which the name Lietor Alluvium is proposed. This unit is at present undergoing trenching.

A more complex Quaternary succession has been reported from the central and southern parts of the Province of Alicante[3] but secure ^{14}C dates are still not available for it. Elsewhere in south-eastern Spain the dating of continental deposits is based largely on their relationship to fossil beaches whose age is in dispute or on correlation with other areas[4,5]. Four ^{14}C dates were obtained during the present study (Table 1). The first was on charcoal from a slightly disturbed hearth near the base of the Elche Formation south of Molinicos (JCD 2). It is worth adding that at Ayna the cemented scree yielded a single Palaeolithic artefact, and that concentrations of Mesolithic material were found on the surface of the Elche Formation at various localities near La Dehesa. The Lietor Alluvium supplied datable charcoal in a gravel pit near Liétor (DW 11), in the Rambla de Dilar, where it forms a terrace remnant (LT 1), and south of Molinicos, where it is represented by colluvium overlying the Elche Formation (JCD 1). The three dates tally satisfactorily, and indicate an age of about 700 yr for the mid-point in aggradation.

A threefold post-Tertiary valley sequence similar to that outlined here has been reported from numerous valleys in other parts of the Mediterranean Basin and in the Near East[6,7] and also in central Périgord[8]. It has been suggested[6] that the tufa reflects greater spring discharges than those of today, that the older of the alluvial fills points to active frost weathering in the uplands coupled with the incidence of intense but short-lived and infrequent rains comparable to those now prevailing at lower latitudes, and that accumulation of the younger fill was in response to a temporary southward shift in the depression belts of Europe. The onset of the last of the three episodes

Table 1 ^{14}C dating of rock units

Rock unit	Sample no.	Laboratory no.	Grid reference	Location	Age (yr b.p.)	$^{13}C \cdot {}^{12}C$ rel. PDB (‰)	Material	Remarks
Elche Formation	JCD 2	SRR-730	7274 × 4317	2°14′22″W 38°27′30″N	39,735 + 1,190 − 1,035	−22.2	Charcoal	At depth of 9 m in 10-m section
Lietor Alluvium	LT 1	SRR-726	7453 × 4363	20°0′46″W	819 ± 35	−23.2	Charcoal	At depth of 1.5 m in 2.5-m section
		SRR-727		38°30′39″N	853 ± 50	−23.4	Charcoal	Repeat analysis
Lietor Alluvium	DW 11	SRR-728	7492 × 4439	1°58′12″W 38°34′43″N	780 ± 110	−24.7	Charcoal	At depth of 1.5 m in 2.5-m section
Lietor Alluvium	JCD 1	SRR-729	7274 × 4318	2°14′22″W 38°27′30″N	419 ± 65	−25.0	Charcoal	At depth of 1 m in 3-m section

varies by up to 500 yr according to location. This has been cited in support of the view that aggradation stemmed from man-induced soil erosion[9], but a preliminary plot against latitude of the ^{14}C dates then available for the deposit indicated a latitudinal decrease in its age which seemed to favour the 'cyclonic' explanation[10].

Figure 2 extends this approach to the other two units. It incorporates the Spanish material and all the other ^{14}C dates whose attribution is certain, as well as limiting archaeological ages derived from historical sites or from excavated material locally calibrated with the help of radiocarbon dating. Granted that the number of points remains small, that they include some (notably those on tufa) which are open to challenge, and that no correction has been made for altitude or other possible

Fig. 2 Plot of age against latitude and provenance of ^{14}C dates for the three deposits reviewed. The upper stippled zone includes the dates for the younger alluvial fill, the lower stippled zone those for the older fill, the blank area between them those for the intervening erosional phase, and the blank triangle at the base those for the tufa. The only dates which do not conform to this scheme are the two from the Dordogne (⊡) within the lower stippled zone, as they are on tufa, but this tallies with field evidence for overlap between tufaceous and alluvial deposition in that area[8]. Archaeological ages (—) include some that immediately predate (⊥) or postdate (⊤) the younger fill. Data from this article and from refs 6–9 and 12.

distorting factors, there is some indication of a northward lag in the transition from tufaceous to alluvial deposition, and of a southward lag in the onset of the second alluvial phase. The graphs also suggest that in both instances aggradation terminated more abruptly than it began, just as the trenching of individual valley fills often progresses more rapidly than does their accumulation[11]. Figure 2 could be taken to imply that the climatic conditions responsible for the older of the alluvial fills emanated from lower latitudes, whereas those that produced the younger fills stemmed from higher latitudes, both inferences being in accord with the interpretations prompted by the field evidence. It remains to be seen whether additonal ^{14}C dates will support the graphs so that their gradients can be used to test palaeo-climatic hypotheses with some confidence.

C. VITA-FINZI

*University College London,
Gower Street,
London WC1E 6BT, UK*

1 Ruddiman, W. F., and McIntyre, A., *Quat. Res.*, 3, 117–130 (1973).
2 López Bermúdez, F., *La Vega Alta del Segura* (University of Murcia, Murcia, 1973).
3 Cuenca Payá, A., and Walker, M. J., *Actas I Reunión Nac. Grupo Trab. Cuaternario*, 2 (edit. by Aleixandre, T., *et al.*), 15–38 (Inst. Lucas Mallada, Madrid, 1975).
4 Solé Sabarís, L., *Quaternaria*, 6, 309–342 (1962).
5 Gigout, M., *C.r.hebd. Séanc. Acad. Sci., Paris*, 249D, 1774–1776 (1959).
6 Vita-Finzi, C., *The Mediterranean Valleys* (Cambridge University Press, Cambridge, 1969); *The Mediterranean Sea* (edit. by Stanley, D. J.), 43–46 (Dowden, Hutchinson and Ross, Stroudsburg, 1972); *Georgrl J.*, 139, 414–421 (1973); *ibid.*, 141, 415–420 (1975).
7 Bintliff, J. L., *Proc. prehist. Soc.*, 41, 78–84 (1975).
8 Vita-Finzi, C., *Nature*, 250, 568–570 (1974).
9 Butzer, K. W., *Perspectives on Environment* (edit. by Manners, I. R., and Mikesell, M. W.), 57–78 (Ass. Am. Geog., Washington, 1974).
10 Vita-Finzi, C., *Recent Earth History* (Macmillan, London, 1973).
11 Leopold, L. B., Wolman, M. G., and Miller, J. P., *Fluvial Processes in Geomorphology* (Freeman, San Francisco, 1964).
12 Choubert, G., Faure-Muret, A., and Maarlveld, G. C., *C.r.hebd. Séanc. Acad. Sci., Paris*, 264D, 434–437 (1967).
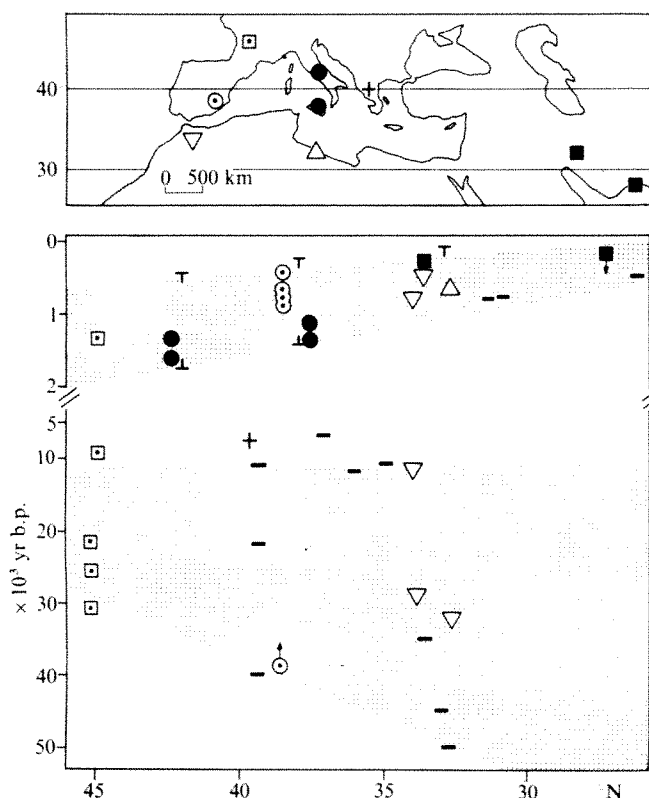
Effect of mineral precipitation on isotopic composition and ^{14}C dating of groundwater

I PRESENT here a method for determining the combined effects of concurrent mineral precipitation and dissolution on the carbon isotope composition of groundwater. The results are applied to the 'adjustment problem' in the ^{14}C dating of groundwater, and are illustrated using published data from the London Basin chalk aquifer. Some ^{14}C ages are changed by more than 10,000 yr.

'Raw' ^{14}C data from groundwaters must be adjusted to allow for dilution attributable to the dissolution of bedrock

carbonate which contains no ^{14}C. Either stable isotope measurements or conventional geochemical measurements may be used to determine the adjustment factor. In both cases 'initial' (that is, at the start of conditions which are closed to carbon dioxide) values of either ^{13}C content (δ^{13}C) or total dissolved inorganic carbon molality (mC) must be known[1]. These initial values are often difficult to estimate. In addition, even adjusted ages are subject to errors arising from various forms of contamination within the aquifer[2]. Nevertheless, provided the aquifer geochemistry is well understood, reasonably reliable ages can be determined.

One important and relatively common process has either not been considered, or been considered inadequately, in past studies of carbon isotopes in groundwater: the concurrent dissolution and precipitation of two minerals (that is, incongruent dissolution, if both are carbonate minerals). Because of isotopic fractionation between aqueous phase and solid carbon-containing species, the isotopic composition of any precipitate will differ from that of the parent solution. For dissolution, the dissolved material will generally differ isotopically from the solute. Thus, both precipitation and dissolution must, in general, alter both the stable and unstable isotope composition of a groundwater. This paper presents the first quantitative estimate of the combined effects of these processes.

Let mD be the contribution to changes in mC over a flow-path attributable to dissolution of one mineral, and let mP be the contribution attributable to precipitation of another mineral. Over a flow-path increment the change in mass of ^{13}C carried by the groundwater must balance the net gain attributable to dissolution and loss attributable to precipitation. This mass balance may be expressed mathematically as

$$d(mC\delta^{13}C) \approx \delta^{13}C_s d(mD) - \delta^{13}C_p d(mP)$$

where 's' and 'p' denote bedrock and precipitate values. In terms of the fractionation factor between precipitate and solution

$$F = \delta^{13}C_p - \delta^{13}C$$

this becomes

$$mCd(\delta^{13}C) + \delta^{13}Cd(mC) \approx \delta^{13}C_s d(mD) - (\delta^{13}C + F)d(mP)$$

If $\alpha = d(mD)/d(mP)$ is the relative dissolution to precipitation contribution to mC, then, for $\alpha =$ constant and $F =$ constant, and noting that $d(mC) = d(mD) - d(mP)$, the above can be integrated to give

$$\delta^{13}C - \delta^{13}C_s \approx X(\delta^{13}C_0 - \delta^{13}C_s) + F(X-1)/\alpha \qquad (1)$$

where

$$X = \left\{ \frac{mC}{mC_0} \right\}^{\alpha/(1-\alpha)} , \alpha \neq 1$$

$$X = e^{-mP/mC_0} , \alpha = 1$$

here, '0' denotes the value at the commencement of concurrent dissolution and precipitation. This would normally be preceded by a dissolution-only period. In equation (1), $X(\delta^{13}C_0 - \delta^{13}C_s)$ is the contribution attributable to precipitation and dissolution ignoring fractionation, and $F(X-1)/\alpha$ is the fractionation effect. As α tends to infinity (that is, dissolution dominates) this term becomes small and the usual dissolution-only result obtains.

For ^{14}C, the corresponding relationship is

$$A \approx A_0 X \qquad (2)$$
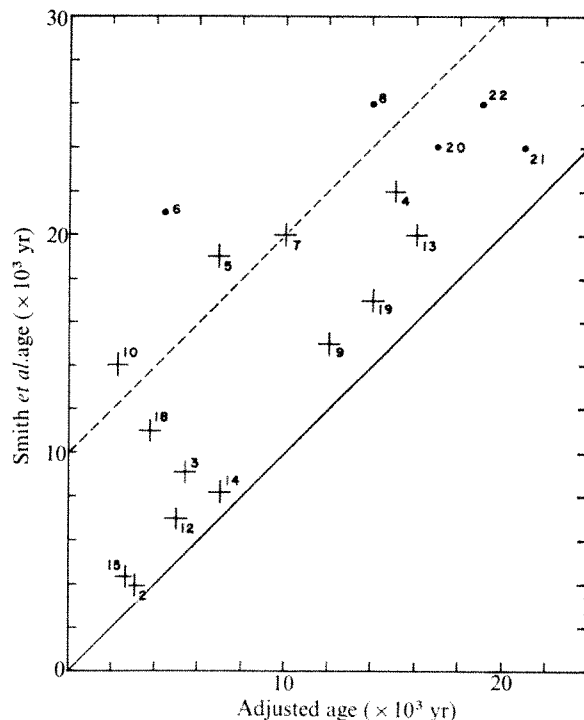
Fig. 1 Comparison of groundwater ages before and after adjustment for the effects of concurrent mineral precipitation and dissolution. Numbers, site numbers used by Smith *et al.*[3,4]; solid circles, data values regarded by Smith *et al.* as having the greatest error margins. For points to the left of the dashed line the precipitation–dissolution adjustment is greater than 10,000 yr.

where A is the activity (in % modern). Detailed derivations of these results will be given elsewhere.

These results are based on the assumptions of constant α and F. Constant α implies constancy of the relative driving forces for dissolution and precipitation. Although made here in the interest of mathematical tractability, I believe that, in the absence of detailed geochemical information from particular aquifers, it is the best possible assumption. The fractionation factor, F, varies considerably with pH and temperature; however, at the relatively high pH values needed to allow concurrent dissolution and precipitation, F is largely independent of these variables and is approximately 2.5‰.

After extensive combined dissolution and precipitation (that is, X tending to zero), equations (1) and (2) show that the isotope concentrations tend to limiting values of $\delta^{13}C_s - (F/\alpha)$ and 0%. Since $F \approx 2.5$‰ this allows $\delta^{13}C$ to approach the bedrock value closely. In addition, ^{14}C activity may become very low independently of groundwater age. As an example, if $\alpha = 1$, $\delta^{13}C_s = 0$‰, and $F = 2.5$‰, the limiting $\delta^{13}C$ is -2.5‰, much higher than expected in the absence of mineral precipitation. This limit corresponds to a steady state where 0‰ carbon is added to the groundwater by dissolution, while 0‰ carbon (that is, the groundwater value plus 2.5‰ fractionation) is being lost at an equal rate by precipitation.

Equation (1) can be used to adjust ^{14}C ages and so give an idea of the magnitude of dissolution–precipitation effects on the ^{14}C dating of groundwaters. Ages adjusted in this way will, necessarily, differ from those obtained using the dissolution-only formula (that is, equation (1) with $F = 0$). The difference will be one half life when $X = F/[F - \alpha(\delta^{13}C_0 - \delta^{13}C_s)]$. Greater differences will occur for smaller values of X (that is, as $\delta^{13}C$ approaches its limiting value). Values of X less than the above correspond to $\delta^{13}C$ values within about 3‰ of the limiting value. Since age errors of one half life (or more) are quite significant, a proper consideration of precipitation effects can be extremely important. As $\delta^{13}C$ approaches its limiting value the adjustment factor becomes increasingly sensitive to small errors in the data and/or the model. Thus, unfortunately,

the uncertainty in the adjustment increases as the magnitude of the adjustment increases.

To illustrate the possible magnitude of this effect I have re-evaluated isotope data from the London Basin chalk aquifer given by Smith et al.[3,4]. These authors state that the anomalously high values of $\delta^{13}C$ observed in this aquifer are most probably attributable to concurrent dissolution and precipitation. In adjusting ^{14}C ages, however, they did not take fractionation effects into account, and used the standard dissolution-only formula. In the London Basin, $\delta^{13}C_s \approx 2.35\%_0$. The chemical evolution of the groundwater seems to be closed-system dissolution (during which $\delta^{13}C$ evolves from a soil-zone value of $-26\%_0$ to around $-13\%_0$) followed by concurrent dissolution and precipitation. The dissolution-only period reduces the ^{14}C activity to around 50%, so that the age may be estimated by

$$t = \frac{5,730}{\log 2} \log \left(\frac{50X}{A_m} \right)$$

where A_m is the measured groundwater activity and X is given (using equation (1) and assuming $\alpha = 1$) by

$$X = \frac{(-\delta^{13}C - 0.15)}{12.85}$$

Figure 1 compares the ages given by Smith et al.[3,4] with ages recalculated using the relationship defined here. The differences are, in some instances very large: more than 10,000 yr in four cases. Adjustments of this magnitude are of considerable importance. In support of these calculations it is reassuring that none of the $\delta^{13}C$ values observed by Smith et al.[3,4] exceeds the limiting value of $-0.15\%_0$ implied by equation (1).

Smith et al.[3] have suggested that some of the London Basin water may be recharge dating from before the last glacial maximum. This suggestion can be re-examined in the light of the revised ages (Fig. 1). After the last glacial maximum, the earliest recharge to the aquifer probably occurred during the first post-glacial warm period around 13,000 yr b.p. Ages in excess of this would confirm the suggestion of Smith et al.[3,4]. Although many ages are substantially younger after readjustment, some of the ages still, indeed, exceed 13,000 yr b.p.; however, in view of the large uncertainties involved in allowing for mineral-precipitation effects, and the proximity of many $\delta^{13}C$ values to the theoretical limiting value (when adjustment uncertainty is greatest), the possibility that all the waters are younger than 13,000 yr b.p. must remain open.

Nevertheless, a direct interpretation of the ages shown in Fig. 1 supports the suggestion of Smith et al.[3,4]. Some of the readjusted groundwater dates do fall within the last glacial maximum period. Since recharge at this time was probably negligible, if these ages are correct then the relevant samples are most probably older waters 'contaminated' by mixing with post-glacial recharge.

In conclusion, mineral precipitation within an aquifer can have a very significant influence on the groundwater isotope geochemistry. Ignoring mineral precipitation, or neglecting isotopic fractionation during precipitation, may lead to estimates of ^{14}C ages which are many thousands of years too old. Stable carbon isotope data may be used to adjust ages, but the largest adjustments are those which are subject to the greatest uncertainty.

I thank R. A. Downing for providing a pre-publication copy of the London Basin results[3,4].

T. M. L. WIGLEY

Climatic Research Unit,
University of East Anglia,
Norwich NR4 7TJ, UK

1 Wigley, T. M. L., Wat. Resour. Res., 11, 324–328 (1975).
2 Pearson, F. J., Jr, and Hanshaw, B. B., in Isotope Hydrology 1970, 271–286 (International Atomic Energy Authority, Vienna, 1970).
3 Smith, D. B., Downing, R. A., Monkhouse, R. A., Otlet, R. L., and Pearson, F. J., Wat. Resour. Res., 12, 392–404 (1976).
4 Smith, D. B., Otlet, R. L., Downing, R. A., Monkhouse, R. A., and Pearson, F. J., Nature, 257, 783–784 (1975).

Calcium oxides of high reactivity

A RECENT study of the kinetics of decomposition of calcite ($CaCO_3$) single crystals in vacuo showed that if the reaction was interrupted before completion a 30-μm layer of a poorly crystalline material was present between the undecomposed $CaCO_3$ and a layer of normal polycrystalline CaO^1. It was hypothesised that the material of this 30-μm layer is a metastable form of CaO that transforms irreversibly to the stable polycrystalline oxide when the accumulated strain exceeds a critical level. If so, the free energy of formation of the metastable oxide from the stable oxide must lie in the range between zero and $+31,500 - 21T$ J mol^{-1} and might well lie at the positive end of this range[2]. Such a metastable oxide should be chemically more reactive than the stable oxide. If the hypothesis is true, then the principal product of decomposition in vacuo of calcite particles of diameter < 30 μm should be the metastable oxide. We present here evidence that this inference is correct, and we demonstrate that not only this oxide, but also a highly crystalline oxide which is produced when large calcite crystals are decomposed in vacuo, reacts more vigorously with water than does the product of calcite decomposition in air or nitrogen.

Calcite ground to a powder formed of block-like rather sharp-edged particles with cross sections ranging from 1 to 10 μm (Fig. 1) was decomposed in air or nitrogen in the range of temperatures between 650 °C and 800 °C. The product was composed of particles which had rounded surfaces and neck areas which indicate considerable sintering (Fig. 2). The product gives the diffraction pattern of highly crystalline CaO.

When the same $CaCO_3$ powder was decomposed in vacuo in the range of temperatures between 450 and 750 °C, there was little change in the apparent particle shape from those of the starting material, and little sintering (Fig. 3). It can be calculated from the relative molar volumes of $CaCO_3$ and CaO that the powder particles produced in vacuum decomposition, like the product of $CaCO_3$ single crystals in vacuo[1], must have $> 50\%$ porosity. Unlike the powders decomposed in air or nitrogen, the powders decomposed in vacuo were poorly crystalline, but did show weak diffraction peaks of the normal (NaCl type) CaO structure (Fig. 4). Heights of the (100) peak showed no

Fig. 1 Scanning electron micrograph of $CaCO_3$ powders.

I μm
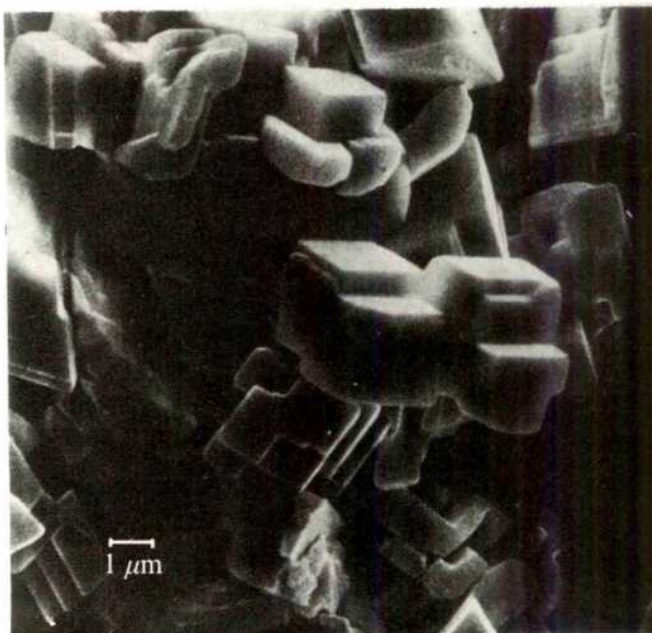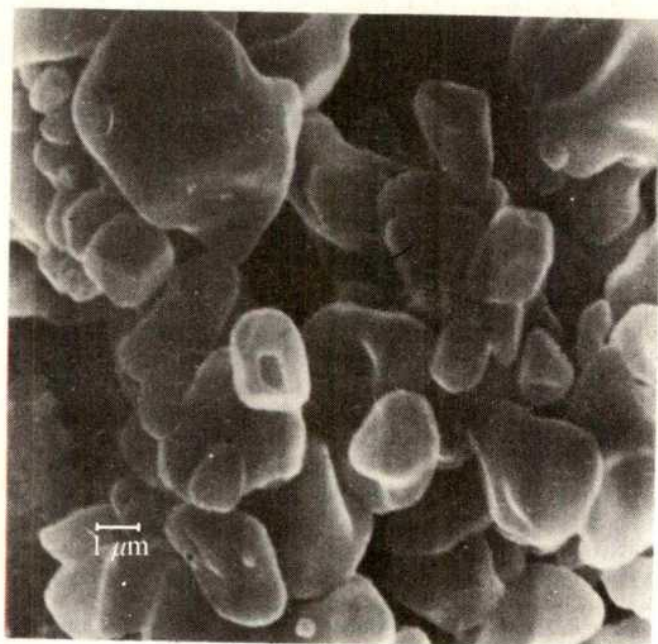
Fig. 2 Scanning electron micrograph of CaO produced from decomposition of $CaCO_3$ powders in dry nitrogen.

clear systematic dependence on time or temperature of heating (as high as 1,050 °C).

The oxide produced by decomposition of $CaCO_3$ powder in air or nitrogen showed no measurable $Ca(OH)_2$ diffraction peaks when ground for 10 to 15 min in air and gained only 1% in weight when exposed to an atmosphere of water vapour at 140 °C. In contrast, the poorly crystalline oxide produced by decomposition of the same powder *in vacuo* is completely converted to poorly crystalline $Ca(OH)_2$ by grinding in air for 10 min or by 3-min exposure to the water vapour stream at 140 °C. The poorly crystalline hydroxide in turn transforms exothermally to the crystalline hydroxide when heated in air to ~290 °C.

Surprisingly, the product of complete decomposition of single crystals of calcite *in vacuo*, even though a CaO with a degree of crystallinity little different from that of the CaO produced by

Fig. 3 Scanning electron micrograph of CaO produced from decomposition of $CaCO_3$ powder *in vacuo*.
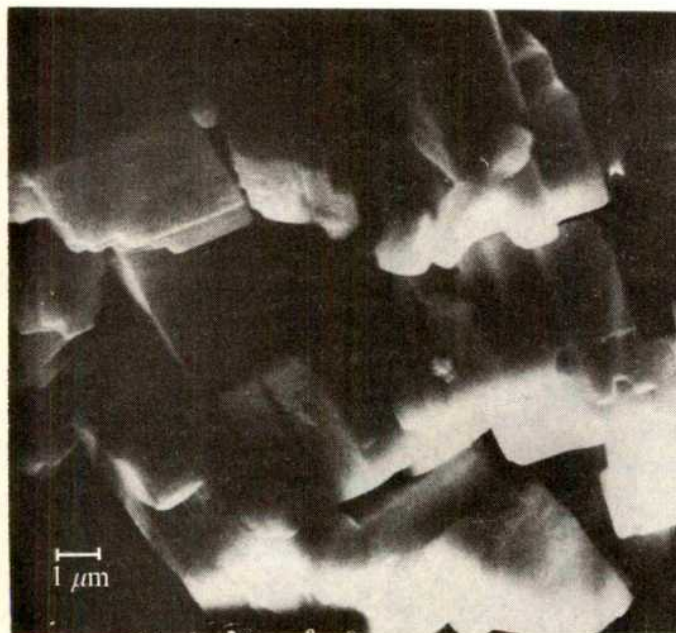
calcite decomposition in air, reacted with water vapour at rates that were not distinguishable by our hydration tests from rates for the poorly crystalline product of vacuum decomposition of $CaCO_3$ powders. This observation shows that the high rates of reactivity of the oxides prepared *in vacuo* are primarily consequences of the high internal surface areas of the block-like products of vacuum decomposition. The blocks produced *in vacuo* have >50% porosity[1], yet the pores are too small to observe at ×30,000. Pore diameters probably average well under 0.01 μm so that solid state diffusion over only short distances is necessary for the hydration reaction. The rounded forms of the oxide particles prepared by decomposition in air must be a consequence of extensive condensed phase diffusion, a process that would probably close the intra-particle pores.

In line with our original hypothesis that the 30-μm layer observed in the interrupted decomposition of a calcite single crystal *in vacuo* is a metastable form of CaO that transforms under sufficient strain to the normal crystalline form, the oxide
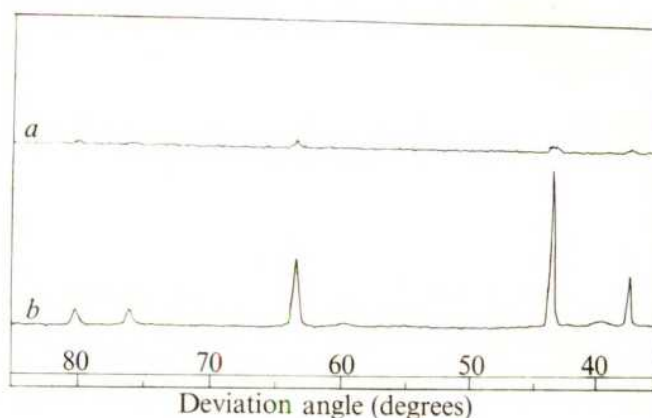
Fig. 4 X-ray diffractometer trace (cobalt Kα radiation). *a*, CaO produced from decomposition of $CaCO_3$ *in vacuo*; *b*, CaO produced from decomposition of $CaCO_3$ powders in air.

of this layer is more reactive towards moisture than the other forms here identified. The X-ray peak which in the earlier study[1] was presumed to be a reflection for a metastable oxide is instead the strongest peak of the poorly crystalline hydroxide. The 30-μm oxide layer, after exposure to room temperature air for <1 h, shows only the diffraction peaks for the amorphous hydroxide. In the same conditions, the other products of *in vacuo* decomposition, if not ground, give patterns of either well crystallised or poorly crystalline CaO, as reported above.

DARIO BERUTO

Laboratorio di Chimica,
Facoltà di Ingegneria,
Università di Genova,
Italy

ALAN W. SEARCY

Materials and Molecular Research Division,
Lawrence Berkeley Laboratory, and
Department of Materials Science and Engineering,
University of California,
Berkeley, California 94720

1 Beruto, D., and Searcy, A. W., *J. chem. Soc., Faraday Trans.*, *I*, **70**, 2145–53 (1974).
2 Searcy, A. W., and Beruto, D., *J. phys. Chem.*, **80**, 425–9 (1976).

Wave power availability in the NE Atlantic

FOLLOWING Salter's[1] proposals for the extraction of energy from sea waves, we are working on the prediction of wave power output from devices situated at favourable coastal sites around the UK. While it is hoped that adequate spectral data from sites closer inshore will become available within the next year or so, the best present data for the North-east Atlantic are from O.W.S. India[2]. We describe here the power available and, on simplified assumptions, the amount that might be extracted by ducks of various diameters. (The Salter 'duck' is a rocking cam-shaped device which is designed to give a high efficiency of energy extraction over a wide frequency band[3].) Confirmation of the validity of the important assumption of additivity of power outputs from wave components of different periods is provided in a separate paper[4].

The power, P, in a sinusoidal wave train in deep water is kTH_{rms}^2 per unit width of wave front, where H_{rms} is the root mean square displacement (that is, the standard deviation) of the water surface, T is the wave period and $k = \rho g^2/4\pi \simeq 7.82$ kW m^{-3} s^{-1}. Thus for a mixed wave train of spectrum dS ($H_{rms}^2 = \int dS$), on linear theory

$$P = k\int T dS = kT_e H_{rms}^2$$

where T_e is called the energy period.

The seasonal joint distributions of T_e and H_{rms} at O.W.S. India are shown in Fig. 1, with contours of P. They are calculated from Hoffman's set of 307 spectra[2] which were selected from a random sample of just over a thousand 10–15 min wave records, taken over the period 1954–67; we weighted each spectrum to restore the correct frequency of occurrence of each band of windspeed in each season. The overall average power is 91 kW m^{-1}.

The predictions given here are for ducks operating with a simple power output limit and (as in the laboratory) relative to a fixed axis. More realistic physical limits on the torque and angular displacement are important for the estimation of the most economic size and characteristics for full-scale ducks, but should not much affect the range of predicted outputs. The related problems of backbone (axis) movement and the directional wave spectrum are more serious; accurate predictions must await laboratory experiments using a less constrained mounting and data on directional spectra at likely wave power sites.

Figure 2 shows the distribution of power by frequency. (For clarity, autumn and spring are omitted; the former would be close to, the latter rather below, the whole year histogram.) Superimposed are efficiency curves for ducks of diameter $d = 6$, 10 and 16 m, scaled up from the experimental curve (for $d = 10$ cm) of Salter, Jeffrey and Taylor[3]. This curve is the result of attempts to optimise performance at low frequencies, so as to minimise the size of duck required for a given sea. The falling off of efficiency at low frequencies is not unexpected, since it cor-

Fig. 1 Frequency of occurrence at O.W.S. India in ‰ of the possible combinations of T_e and H_{rms}, with contours of P in kW m^{-1}. *a*, Winter (Dec-Feb); *b*, spring (Mar-May); *c*, summer (Jun-Aug); *d*, autumn (Sep-Nov).
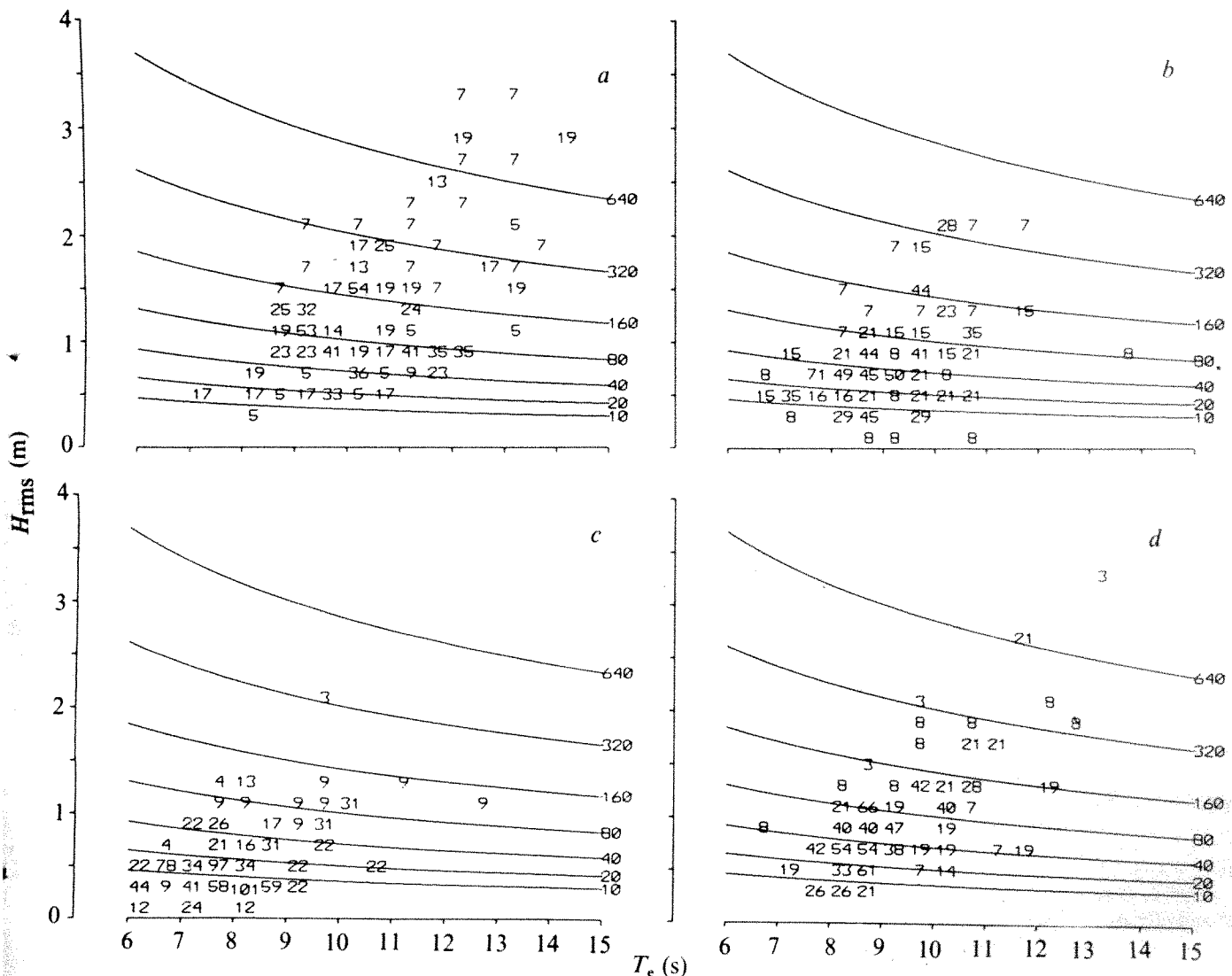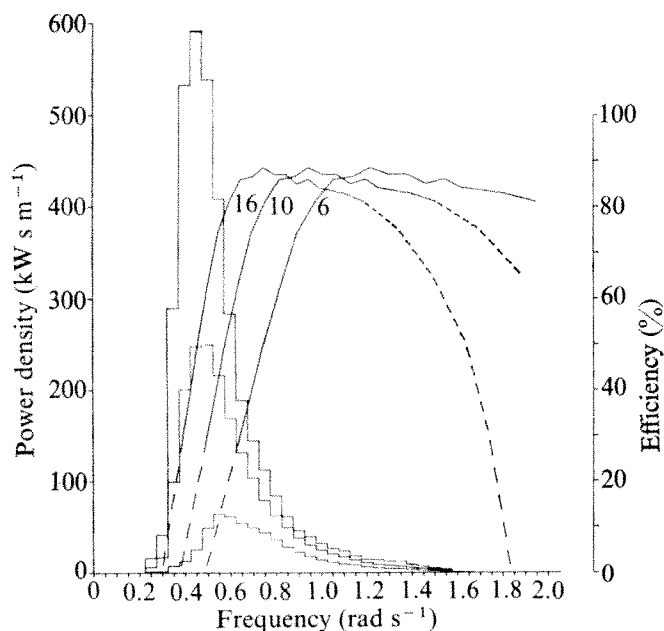
Fig. 2 Distribution of power by frequency for the whole year (middle histogram), winter (top) and summer (bottom); together with predicted efficiency curves for ducks of 6, 10 and 16 m diameter (dashed lines indicate extrapolation outside experimental range).

responds quite closely to the proportion of energy travelling in the water layer above depth d ($= 1 - \exp\{-2d\omega^2/g\}$).

The average output for a particular duck size and season, if no power limits are imposed, can then be found simply by multiplying the appropriate two curves in Fig. 2 and integrating. This is the power transferred from sea to duck; the conversion to electricity and transmission to a distant load-centre in a hydrodynamically underprivileged area, such as South-east England, can be done with presently available technology with an efficiency of $\sim 60\%$. To investigate the variability of power levels, it is of course necessary to refer to the individual spectra. The proportion of time for which power outputs might exceed any given level is shown in Fig. 3, with the total power available for comparison.

We have investigated the effect on power output calculations of replacing the real sea spectra by two-parameter (H_{rms} and T_e) spectra of standard shape, the shape chosen being that of the Pierson–Moskowitz spectrum[5]

$$\int_0^\omega dS = H_{rms}^2 \exp\{-b(\omega T_e)^{-4}\}$$

where $b \simeq 1,052$ (dimensionless).

This is not only convenient for present computational purposes, but should also greatly facilitate predictions of the effects of physical and torque overloads and the directionality of real seas. This replacement has little effect on the distributions of power input (see Fig. 3, dashed lines), though estimates for individual spectra have s.e. 4.55%. The largest effects on the predicted average output over the range of diameters 6–18 m are -0.5 kW m^{-1} (for $d = 6$ m in the summer) and $+1.5$ kW m^{-1} (for $d = 12$ m in the winter) which may be attributed to real spectra being respectively broader in calm, and narrower in rough conditions than the standard spectra. We conclude that, while further checks are advisable, such standard spectra appear sufficiently accurate for our present approximate predictive purposes. (When directional data become available we intend to extend this investigation to see whether the addition of one further parameter to allow for directional spread yields a similarly adequate description for predicting behaviour of free-floating duck strings in directionally mixed seas.)

Directional problems apart, there are three limitations on output which we have not considered so far: physical (swept-volume), torque and transmission limits. Because of the high costs of transmission lines[6] the last of these is probably the most critical, though swept-volume overload limits will be important for ducks of diameter < 10 m.

The mean power available when simple power output limits

Fig. 3 The proportion of time for which each power level is exceeded for: *a*, whole year; *b*, winter and *c*, summer. In each case the top curve refers to the total power available; the other solid curves, in descending order, to predicted power output from ducks of 16, 10 and 6 m diameter. The dashed curves (verticals omitted) show the effect, for each of these three cases, of replacing real spectral data by two-parameter standard spectra (see text).
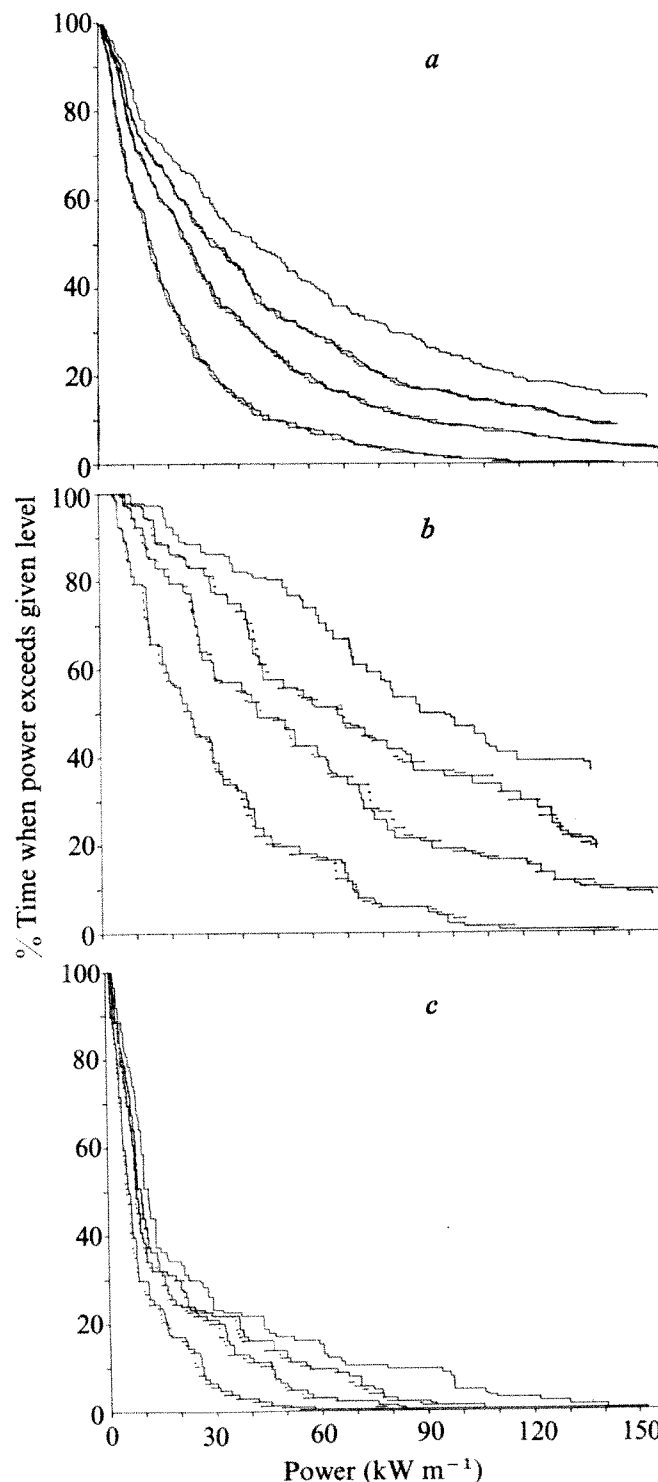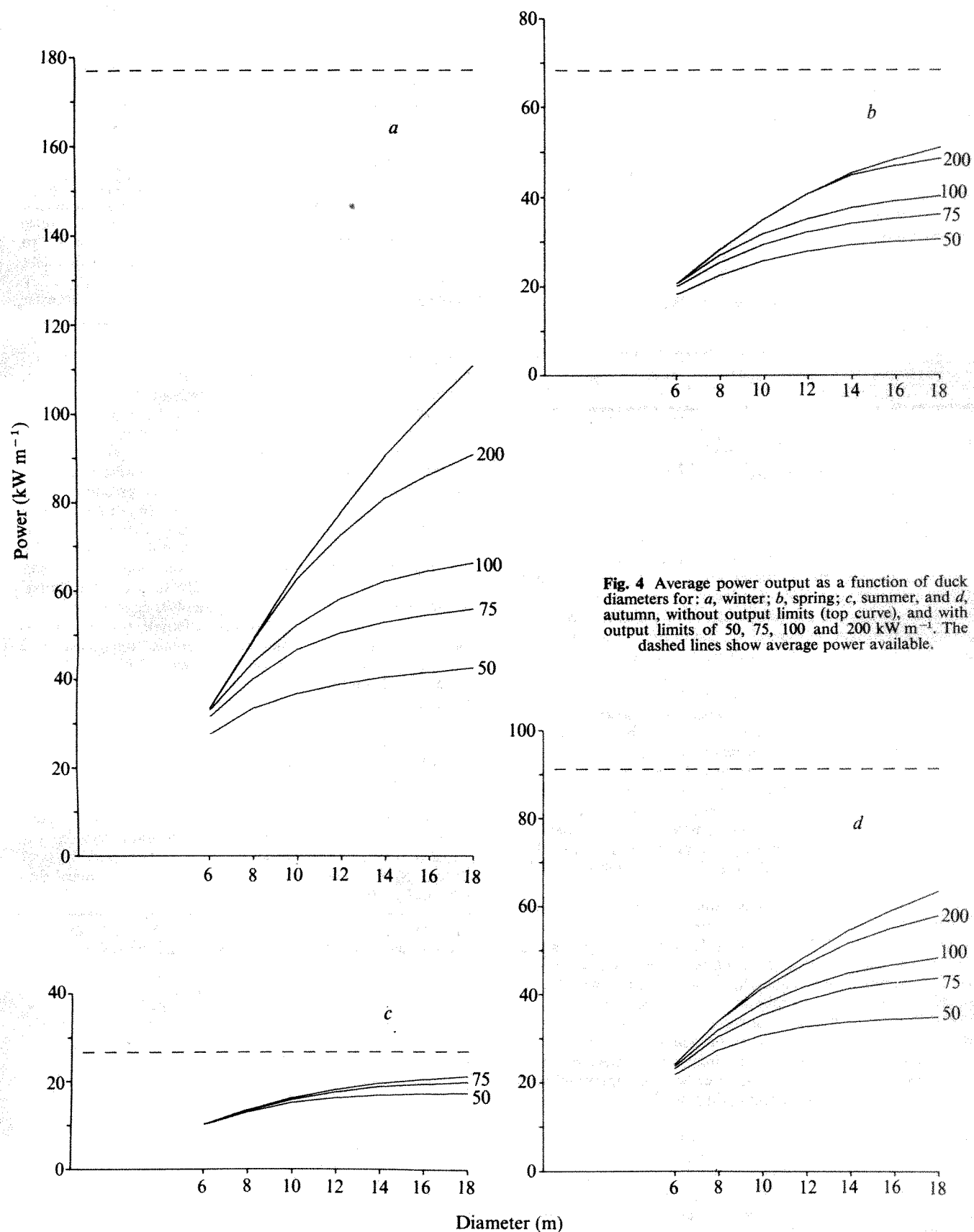
Fig. 4 Average power output as a function of duck diameters for: *a*, winter; *b*, spring; *c*, summer, and *d*, autumn, without output limits (top curve), and with output limits of 50, 75, 100 and 200 kW m^{-1}. The dashed lines show average power available.

are imposed is easily calculable as the area under the relevant curve of Fig. 3 to the left of the desired limit. This is plotted as a function of duck diameter in Fig. 4. As might be expected, the effect of such limits is more severe in winter. Even then the

lowest (50 kW m^{-1}) limit considered gives a substantially higher load factor ($\sim 80\%$ for $d \geq 12$ m), a further reason for preferring a relatively low rated transmission line, rather than going for maximum average output.

This work was supported by the SRC and the Department of Industry.

DENIS MOLLISON

Department of Actuarial Mathematics and Statistics,
Heriot-Watt University,
Edinburgh, UK

O. P. BUNEMAN

Department of Computer Science,
University of Pennsylvania,
Philadelphia, Pennsylvania 19104

S. H. SALTER

Department of Mechanical Engineering,
University of Edinburgh,
Edinburgh, UK

1 Salter, S. H., *Nature*, 249, 720–724 (1974).
2 Hoffman, D., *Sampling of Ocean Wave Spectra at North Atlantic Station India* (SNAME, New York).
3 Salter, S. H., Jeffrey, D. C., and Taylor, J. R. M., *The Naval Architect*, 1, 21–24 (1976).
4 Mollison, D., Salter, S. H., and Taylor, J. R. M., *Edinburgh Wave Power Project Report No. 32* (1976).
5 Pierson, W. J., and Moskowitz, L., *J. geophys. Res.*, 69, 5181–5190 (1964).
6 Whittington, H. W., *Edinburgh Wave Power Project Report No. 33* (1976).

Male emigration and female transfer in wild mountain gorilla

IN many group-living primates, males leave their natal group and transfer to other groups far more commonly than do females (for example, refs 1–4). For only one species—the chimpanzee—is there good evidence that females leave their natal group and transfer more commonly than do males[5,6], but little information has been available on the movements in and out of gorilla groups. The structure of the gorilla population in the Virunga Volcanoes of Zaire and Rwanda[7,8] and elsewhere[9] indicates that males at least leave their natal group: there are a number of lone males; and many groups contain only one fully adult male, but a number of adult females. Schaller[9] implied that males transferred between groups more often than did females. The few data in ref. 10 indicate, however, that only females transfer between groups, although both males and females leave groups. Incorporating data for 8.5 yr from the Karisoke Research Centre in the Virunga Volcanoes (ref. 10 and unpublished results), this paper provides the first extensive evidence that both male and female gorillas leave their natal group, although some sons may remain; that males that leave initially travel alone; and that only females transfer.

Groups and lone males will be called 'units'. A distinction must be made between short 'visits' by animals to other units and longer 'transfers'. A 'transfer' is defined as a stay of more than 1 d with the new unit. Males mature when about 11 yr old, their first recorded emigrations being at about this age (Table 1a). Males that did not emigrate were the leading males or animals that were probably some of their sons (Table 1b). Leader males tended to inhibit young males from mating, and competition for oestrous females may be a cause of male emigration. Young males that meet stiff competition will probably leave; those that do not will probably stay, and these are likely to be the sons of ageing leaders. One such male took over leadership of the group when his father died. Such inheritance of leadership may be a pattern common to most groups.

Females become mature at about 8 yr old, and also tended to leave the group at about the age of maturity; they tended to leave before giving birth (Table 2a), as has been reported for chimpanzees[6]. In contrast to males, it seems that nearly all females eventually leave their natal group. Because females transfer between groups, however, for only the four youngest females in Table 2 can we be sure that the group in which they were first seen was their natal group.

The one female that remained in her natal group (Table 2b)

produced an offspring, probably sired by her own father, or by a male who was probably her half-brother. No multi-parous females with living offspring emigrated from the group in which they were first observed (Table 2b).

Four males were intermittently observed for 2–5 yr after they first emigrated out of the study groups. None transferred and all spent most of their time alone. Also, no male transferred into any of the study groups. It seems that males that leave their natal group become successful breeders, not by transferring and not by ousting established group leaders (as do langur males[11,12]), but by attracting more than one female to form a new group. A male and single female seem not to be a viable breeding unit.

Females that emigrate transfer almost immediately to other units (as reported for chimpanzees by Kawanaka and Nishida[5]), and they tend to transfer more than once (Table 2a). The length of stay with one unit before the next transfer ranged from 3 d to 3 yr 5 months, with a median of 3 months ($n = 13$). Initial transfers were to units with ranges overlapping those of the group in which the female was first observed, but subsequent transfers took some of the females out of touch of their initial group's range. Factors that determined whether or not a female stayed with her new unit will be considered later in this paper. Females transferred about equally often to lone males ($n = 11$ transfers) as to groups ($n = 9$ transfers). A number of these transfers were returns to previous units, two of them (possibly 4) to the presumed natal group. If returns are omitted, more females (6 compared with 4) transferred more frequently (9 transfers compared with 4) to more lone males ($n = 4$, possibly 5) than to groups ($n = 2$).

In only three of more than 20 observed potential transfer situations was a male observed actively to prevent a female from approaching the new male, in spite of its being presumably disadvantageous for a male to lose a female that is not his daughter. It seems that males inhibit female emigration more by prevention of proximity of other males than by herding females: of 12 interactions between units during which females transferred, 10 involved intense displays, and sometimes fights, between the males.

It is presumably advantageous to a male, especially a lone male, to gain an unrelated female. Thus, kidnap of females might be expected (Fossey, p. 334 in ref. 13). Out of eight observed visits or transfers, however, only one was possibly caused by coercion by the new male; the others were clearly of the females' own volition. (The discrepancy between the number of interactions (12) and the number of observed visits or transfers (8) is because much of the course of an interaction can be judged from the spoor alone.)

Table 1 Data on males aged 11 yr or more at the end of the study that (a) emigrated from their probable natal group* or (b) remained in it

a	When first observed		When first emigrated		
n	Age	Class	Age	Class	Transfer
2	>12	M	>14	M	No/No?
1	>11	M	>12.5	M	No
1	9.5	Ad	11.5	M	No
1	8	Ad	12	M	No
†1	5.5	J	10	Ad	No?
‡1	2.5	I	11	Ad	No

b	When first observed		When last recorded		
n	Age	Class	Age	Class	
2	>12	ML	>20	ML	
1	>12	M	>20	ML	
†1	5.5	J	13.5	ML	
†1	3	I	11.5	M	

*Because no male transferred between groups.
ML, Mature leader male; M, mature male (>11 yr); Ad, male aged 8–11 yr; J, juvenile (3–6 yr); I, infant (0–3 yr).
†Almost certainly a son of a leader male; ‡Almost certainly not a son of a leader male. Ages are to nearest 0.5 yr.

Table 2 Data on females aged 8 yr or more at the end of the study that (*a*) left the group in which they were first seen or (*b*) remained in it*

a	When first observed		When first emigrated			
n	Age	Class	Age	Class	Parous	Transfer
2	6.5–7	sAd	11, 13	Ad	Yes	Yes
2	6–6.5	sAd	9.5, 10	Ad	No	Yes
1	5.5	J	10.5	Ad	No	Yes
3	1–2	I	6.5, 8, 8	sAd	No	Yes

b	When first observed		When last recorded		Minimum number of
n	Age	Class	Age	Class	offspring alive
5	>9	Ad	>17	Ad	1, 2, 2, 3, 3
1	1.5	I	9.5	Ad	1

*This group is almost certainly the natal group of the four infants. Ad, Sexually mature female (> 8 yr); sAd, subadult (6–8 yr); J, juvenile (3–6 yr); I, infant (0–3 yr). Ages are to the nearest 0.5 yr.

What factors influence a female's choice of male, and her decision to stay or transfer again? By analogy with ornithological studies[14], and with some data to support the suggestion, the initial decision might depend on the quality of the male's range. Subsequent decisions to stay or transfer might rest on the female's success at raising her offspring in the unit: two mothers who successfully raised their offspring in a new unit stayed; three who did not, left. The infants of at least two of the latter were killed deliberately during interunit interactions.

Finally, why are gorillas and chimpanzees, perhaps the only primates in which females more frequently transfer between units than do males? Four reasons are suggested. The first two are presented in terms of consequences of remaining in the natal group.

● Macaques are the best documented taxon in which males, but rarely females, transfer between groups[2,3]. Also well documented in this taxon is the phenomenon of offsprings' assumption of agonistic ranks close to that of their mothers', and of females keeping their rank from immaturity into adulthood[15,16]. Even if immature chimpanzee and gorilla females assumed ranks close to their mothers' (and gorilla females probably do not), there would be little advantage in doing so because of lack of competition in which dominance would be an advantage. Thus, to a gorilla or chimpanzee female, there are less advantages to remaining in her natal group than there are to at least a macaque female.

● Conversely, for chimpanzee males, there are advantages to remaining within their natal community because of communal defence of range and acquisition of females[6,17,18].

The final two reasons for female, rather than male, transfer in gorillas and chimpanzees are presented as proximate causes of the observed pattern of, first, emigrations, and second, transfers.

● Although the evidence is diffuse, it seems that young adult males and females tend not to copulate with familiar individuals[19]. Given the current state of transfers, females of many, perhaps most, group-living primate species can copulate with unfamiliar mates in their natal group; chimpanzee males can do the same with females; neither gorilla females, nor males (because the dominant male inhibits young males from mating) can do so; and we find that both young male and female gorillas leave their natal group.

● Resident gorilla and chimpanzee males seem more capable than do resident males of other species, of preventing non-group males from entering the group and mating; gorillas, because of small group size and rarity of oestrus; chimpanzees, because of communal attacks on strange males[17,18]. Thus gorilla males that leave their natal group (no chimpanzee males have been known to do so[5,6]) travel alone until females transfer to them, whereas in other species, the males themselves tend to transfer.

For permission to work in their countries, we thank the Governments of Rwanda and Zaire. For permission to work

A. H. HARCOURT
K. S. STEWART
D. FOSSEY

Sub-Department of Animal Behaviour,
Madingley, Cambridge CB3 8AA, UK, and
Karisoke Research Centre,
BP 105, Ruhengeri, Rwanda

1 Carpenter, C. R., in *Primate Behaviour: Field Studies of Monkeys and Apes* (edit. by DeVore, I.), 250–291 (Holt, Rinehart and Winston, London and New York, 1965).
2 Drickamer, L. C., and Vessey, S. H., *Primates*, 14, 359–368 (1973).
3 Norikoshi, K., and Koyama, N., *Symp. fifth Congr. Int. Primat. Soc.*, 43–61 (Japan Science Press, Tokyo, 1975).
4 Packer, C., *Nature*, 255, 219–220 (1975).
5 Kawanaka, K., and Nishida, T., *Symp. fifth Congr. Int. Primat. Soc.*, 173–186 (Japan, Science Press, Tokyo, 1975).
6 Pusey, A., in *The Behaviour of Great Apes: Perspectives on Human Evolution*, 4 (edit. by Hamburg, D. A., and Goodall, J.), (Staples/W. A. Benjamin, London and New York, 1976).
7 Groom, A. F. G., *Oryx*, 12, 207–215 (1973).
8 Harcourt, A. H., and Groom, A. F. G., *Oryx*, 11, 355–363 (1972).
9 Schaller, G. B., *The Mountain Gorilla: Ecology and Behaviour* (University of Chicago, Chicago and London, 1963).
10 Elliott, R. C., *Folia primat.*, 25, 12–24 (1976).
11 Yoshiba, K., in *Primates: Studies in Adaptation and Variability* (edit. by Jay, P. C.), 217–242 (Holt, Rinehart and Winston, London and New York, 1968).
12 Rudran, R., *Folia primat.*, 19, 166–192 (1973).
13 Hinde, R. A., *Biological Bases of Human Social Behaviour* (McGraw-Hill, London, 1974).
14 Verner, J., and Willson, M. F., *Ecology*, 47, 143–147 (1966).
15 Sade, D. S., in *Social Communication Among Primates*, 99–114 (edit. by Altmann, S. A.), (University of Chicago, Chicago and London, 1967).
16 Koyama, N., *Primates*, 8, 189–216 (1967).
17 Bygott, J. D., in *The Behaviour of Great Apes: Perspectives on Human Evolution*, 4 (edit. by Hamburg, D. A., and Goodall, J.), (Staples/W. A. Benjamin, London and New York, 1976).
18 Wrangham, R. W., in *The Behaviour of Great Apes: Perspectives on Human Evolution*, 4 (edit. by Hamburg, D. A., and Goodall, J.), (Staples/W. A. Benjamin, London and New York, 1976).
19 Bischof, N., in *Biosocial Anthropology* (edit. by Fox, R.), 37–67 (Malaby, London, 1975).

Female African wild dogs emigrate

AMONG mammals, the common mechanism of individual transfer between social groups is male emigration. While studying the African wild dog (*Lycaon pictus* Temminck) over a period of 2 yr, we have recorded four positive cases of female group emigration, three possible cases of single female emigration, and only one possible case of male group emigration. From this we conclude that female emigration is the primary mode of interpack transference in African wild dogs. The only other mammal for which this has been reported is the chimpanzee[1].

The African wild dog is a medium-sized carnivore living and hunting cooperatively in packs[2]. Average pack size during our study was 7.2 adults (computed at 3-month intervals for seven packs, an adult being 1 yr and older). Group social organisation in this species usually consists of one breeding pair per pack and a cooperative system of pup rearing by all pack members[3–7]. According to zoo records[7,8] and observations of very small pups at dens[7,9], the sex ratio is unequal at birth, with 55–59% males. Of 78 dogs 2 yr and older (and therefore sexually mature) that we studied, 69% were, however, males.

Fieldwork began in January 1974, and is continuing. Individual wild dogs are recognised by their distinct coat colour pattern (Fig. 1). For field identifications, a photographic recognition file is carried. Life histories and family relationships are recorded for all known dogs. These data are augmented by Hugo van Lawick's photographic recognition file for 1968–72. Our study area is 4,200 km² of grasslands and 1,000 km² of surrounding woodlands in northern Tanzania's Serengeti National Park and Ngorongoro Conservation Area. This includes the entire known ranges of four wild dog packs and part of the ranges of two additional

packs. Pack ranges cover a minimum of 1,500 km², and overlap about 10–50% with the range of each of the three or four neighbouring packs. Surrounding our study area there are at least seven more packs. We believe the wild dog population may be depressed by disease. Furthermore, shooting within the park as recently as November 1973 has entirely exterminated some packs.

Previous studies tended to concentrate on one pack of wild dogs, with only infrequent sightings of other packs. Changes in pack composition between sightings were therefore usually impossible to interpret. There are, however, two verified instances and one probable case, of female group emigration from the years 1968–71 (J. R. Malcolm and H. van Lawick, personal communication) in the same area as our study area.

Instances of emigration among our six study packs are shown in Fig. 2. In our study, emigration was verified when the following criteria were met: (1) the original pack was seen alive and intact after the disappearance of some of its members; (2) the missing dogs were seen alive after they had left; and (3) the missing dogs joined a new pack and did not return to their original pack. There were four verified cases of group emigration of females, involving 14 individuals. Four of these females emigrated twice.

Individual females disappeared from their packs in three further cases and were never seen again. Two of these females were sisters, who disappeared individually. Because they might have been expected to emigrate together, but did not, there is some doubt about whether they really emigrated. We believe that both probably did emigrate, however, because all pack members seemed to be healthy just before each sister's disappearance, one of their sisters had established herself as reproductively dominant, and both were of an age beyond which no females remain in packs as non-breeders. The one remaining sister of the dominant female who stayed in the pack was also able to breed. She was, however, clearly subordinant, and was allowed only limited access to her pups. Her mate was probably one of the subordinant males. In the third case, a lone daughter of the pack disappeared. She had no sisters which could have accompanied her, and she was the same age as other young females whose emigrations were verified. Each of these three females probably emigrated and died. Mortality of emigrating lone adult females may account significantly for the preponderance of adult males in the population.

During our study, there was one unexplained dis-

appearance of three male siblings, just at the age of sexual maturity. They probably emigrated from their parent pack, but were not seen again.

In February 1976 we witnessed the emigration of two daughters from the Seronera pack to the three males of the Kühme pack. We followed them continuously for 6 d, which included their departure from the parent pack, their journey in search of the Kühme males, and their subsequent joining. Their emigration seemed to be precipitated by the proximity of the Kühme males who were without a female. The emigrating females tracked the scent trail of these males, and one of them repeatedly made contact calls. At the initial meeting of the two groups, there was a high level of aggressive pawing, shoving, and riding-up between the sexes, initiated by the females (Fig. 1). This behaviour continued for the subsequent 2 d of observation. Within 2 min of meeting, however, it became obvious which male and female would form the new breeding pair. One male excluded the other two males with threats, and one female excluded her sister with threats and attacks. Within 90 min this pair was urine-marking together, in the manner of a breeding pair[5]. Behavioural details of the emigration, and a general discussion of the dynamics of African wild dog packs, and the phenomenon of pack splits, will be published elsewhere[5].

The following main features of female emigration have emerged in 2 yr of study : (1) Female emigration occurred when a pack already had one breeding female. (2) Emigrating females never joined a pack that already contained a female. (3) Factors which drove or enticed females from their packs are obscure. The suddenness and rapidity of three transfers suggested an outside stimulus precipitating emigration, probably olfactory or visual evidence of a group of males nearby. There was no evidence that oestrus was the initiating factor, although all females were of the age of sexual maturity, and in one case oestrus and conception occurred soon after the females joined new males. (4) Females did not emigrate further than the adjacent or overlapping ranges of neighbouring packs. This was true in all cases, including the group of four females who roamed 4 months before joining new males. (5) Females usually emigrated in groups, if this was possible.

In the existing ecological conditions and low density of African wild dogs, female emigration has been the regular and predictable pattern of interpack transference. Male emigration has been of lesser importance, with no such transfers verified. In packs with offspring of known or
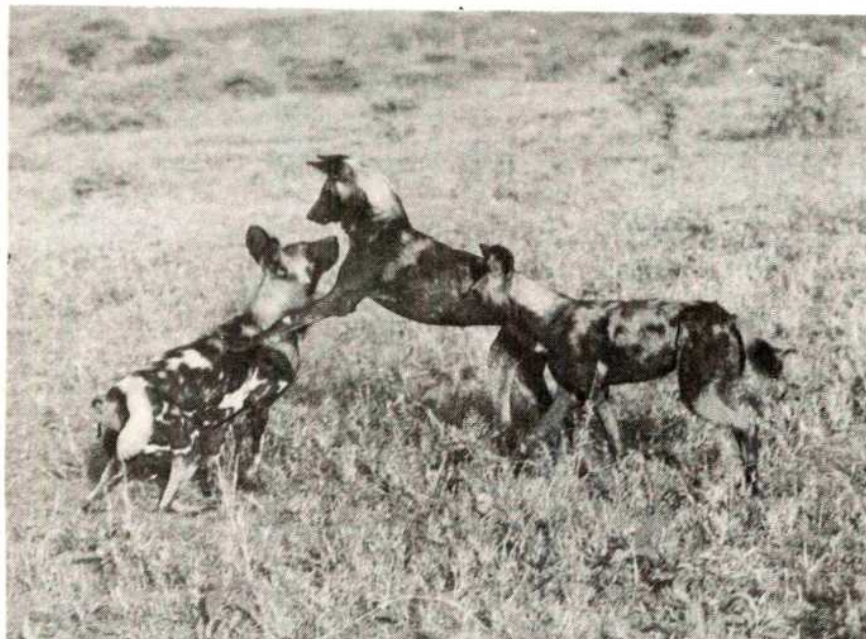
Fig. 1 Side view of three African wild dogs showing distinctive coat patterns used in individual identification. This photograph shows the two sisters (right), who emigrated from the Seronera Pack. During the third morning of their emigration, they met the males of the Kühme pack, who lacked females. A great deal of aggressive pawing, shoving, and riding-up was initiated by the females.

Fig. 2 Diagram of female transference among six study packs of African wild dogs. Three of the four transfers occurred within several days; the fourth took four months. Four females of the Simba pack underwent a secondary emigration after a sister became established as a breeding female in the Seronera Pack. Vertical columns represent the same individual through time. ♂, Adult male; ♀, adult female, ♂♀, dominant breeding pair, ♂ ♀, yearling littermates, ×, dead, ?, disappeared.
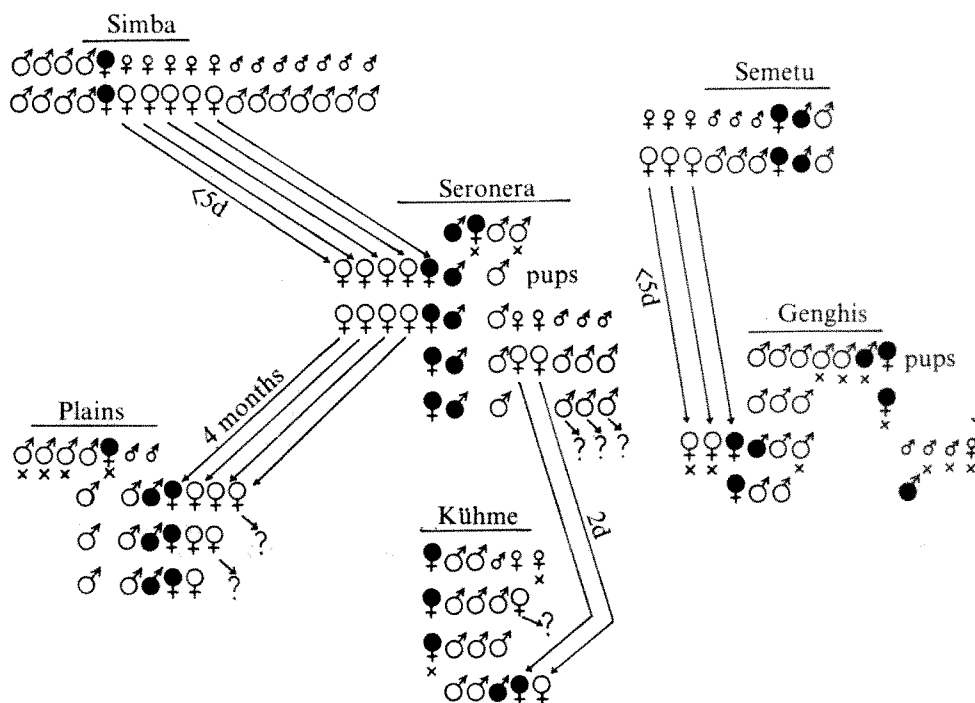
reliably estimated age, all daughters left when they were 18–24 months old. Their male siblings usually stayed in the parent pack. A process of secondary emigration existed, whereby sisters of a new breeding female left her after she attained this status. In combination with probable mechanisms which delay some females' breeding, emigration accounts for the usual occurrence of only one breeding female per pack. Differential mortality of females who emigrate singly or in small groups may account for the sex ratio favouring males more in adults than in immature dogs.

We thank the Trustees and Director of Tanzania National Parks, the Director of the Serengeti Research Institute, the Chief Park Warden of the Serengeti National Park, and the Conservator of Ngorongoro Conservation Area for permission to do this research. Baron Hugo van Lawick and James R. Malcolm shared information from their earlier fieldwork. Many others, especially Jerry Rilling, helped by providing data and photographs of their sightings of the dogs. Financial support was provided by the East African Wildlife Society, Max Planck Institut für Verhaltensphysiologie, the African Wildlife Leadership Foundation, the Shikar-Safari Club, Sigma Xi: The Scientific Research Society of North America, The Explorers Club, The Zoological Society of Philadelphia, the Fauna Preservation Society, Baron Hugo van Lawick, and The Fund for Animals Inc. The manuscript was critically read by J. D. Bygott, R. D. Estes, L. J. Folse, J. J. R. Grimsdell, J. P. Hanby, J. D. and S. H. Ligon, J. R. Malcolm, S. J. McNaughton, J. Rood, H. van Lawick, and W. Wickler.

LORY HERBISON FRAME
GEORGE W. FRAME

Serengeti Research Institute,
Post Office Seronera, via Arusha,
Tanzania

1 Pusey, A., in *The Great Apes* (edit. by Hamburg, D. A., and Goodall, J.), (Holt, Rinehart and Winston, New York, in the press).
2 Malcolm, J. R., and van Lawick, H., *Mammalia*, 39, 231–240 (1975).
3 Kühme, W., *Zeit. f. Tierpsych.*, 22, 495–541 (1965).
4 Estes, R. D., and Goddard, J., *J. Wildl. Mgmt.*, 30, 52–70 (1967).
5 van Lawick, H., (with Goodall, J.), in *Innocent Killers*, 49–101 (Collins, London, 1970).
6 van Lawick, H., *Solo*, (Collins, London, 1973).
7 Schaller, G. B., in *The Serengeti Lion*, 321–344 (University of Chicago Press, Chicago and London, 1972).
8 Kleiman, D. C., and Eisenberg, J. F., *Anim. Behav.*, 21, 637–659 (1973).
9 Frame, L. H., Frame, G. W., Malcolm, J. R., and van Lawick, H., *Zeitschrift für Tierpsychologie* (in the press).

Allometry of neonatal size in eutherian mammals

I HAVE demonstrated that neonatal weight in 15 species of anthropoid primates scales at a power of maternal weight of 0.70 (ref. 1). Gould[2] noted the closeness of this value to the standard metabolic exponent of 0.75, and asked if this similarity was more than coincidental. In an attempt to answer Gould's query, I have investigated the scaling of neonatal size in eutherian mammals and primates in particular.

Incorporated in this study are data on weights of neonates (foetus at term ± 5 d) and adult, non-pregnant females of 36 eutherian species ('mouse to elephant' curve)[3] and 29 primate species[4], in addition to previously unpublished data on 287 neonates and 316 females of seven species of *Macaca* (*M. fascicularis, M. radiata, M. nemestrina, M. niger, M. mulatta, M. arctoides* and *M. fuscata*).

Among eutherian mammals and subordinate eutherian taxa, neonatal and maternal weight are allometrically related through the equation

$$y = bx^a$$

where y is neonatal weight (weight of the whole litter if there is more than one offspring), b the initial size constant (y intercept), x the maternal weight, and a the exponent of allometry (slope)[3,4]. Table 1 and Fig. 1 show that there is a concomitant change of a and taxonomic level. This situation parallels the finding for the brain size–body size relationship[2]. On the other hand, for the scaling of the other principal organs (for example, heart, lungs and liver), the relationship between a and taxonomic level is far less regular[5–8] (Table 1 and Fig. 1).

An assessment of the various values of a with taxonomic level requires a causal explanation. In the scaling of brain size, the combined influence of differences in the evolutionary level of brain development and differences in mean body size of groups at the same taxonomic level accounts for differing values of a (ref. 2). Similarly, in the scaling of neonatal size, differences in the structure and physiology of foetal membranes and placenta in combination with differences in mean body size must be considered. Thus, a meaningful interpretation of a
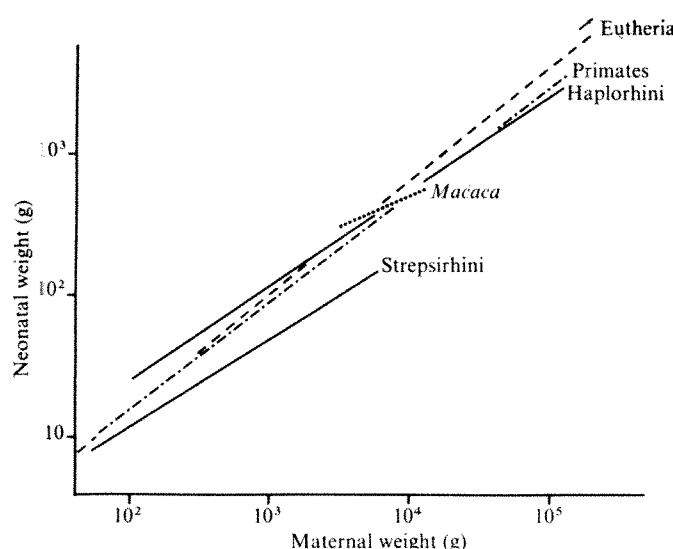
Fig. 1 Neonatal–maternal weight relationships in Eutheria, Primates, Strepsirhini, Haplorhini and *Macaca* (based on values from Table 1).

in the scaling of neonatal size requires that either corrections are made for structural and physiological dissimilarities during foetal development, or comparisons are limited to taxa with members that are structurally and physiologically similar.

Among the Eutheria basic similarities in foetal membrane development only exist during early blastocyst formation[9], after which there is considerable variation in the development of foetal membranes and placenta among the higher taxonomic categories (orders and suborders)[9]. Because it is virtually impossible to quantify the influence of structural and physiological differences in foetal development on neonatal size for a taxon as diverse as Eutheria, a meaningful interpretation of a is difficult. In any case, a (0.83) for Eutheria is clearly above the standard metabolic exponent.

Primates are also variable in the foetal development of many features. For example, there is a strong dichotomy in the structure of the chorio-allantoic placenta and associated foetal membranes. The members of the suborder Strepsirhini have a non-invasive, diffuse, epitheliochorial placenta, a large allantois and a transitory chorio-vitelline placenta, whereas an invasive, discoidal, haemochorial placenta and rudimentary allantois characterise members of the suborder Haplorhini[9]. Although the standard metabolic exponent (0.75) is within a 95% confidence interval of a (0.78 ± 0.05) for Primates, this similarity may be merely coincidental. Actually, the value of a in Primates is due to the combined result of lower but similar values of a in the two suborders, the separation of the two suborders by an allometric shift (differing values b due to differences in foetal membrane and placental structures), and a difference in mean body size between the two suborders.

In contrast to the infraclass Eutheria and the order Primates, structural and functional similarity in foetal membranes and placenta exists within each of the suborders Strepsirhini and Haplorhini. It is interesting that, whereas the standard metabolic exponent (0.75) lies outside the 95% confidence interval of a for both Strepsirhini (0.63 ± 0.05) and Haplorhini (0.69 ± 0.04), another 'standard' exponent (0.67) is within these confidence limits. The value of 0.67 is known as the standard exponent for the interspecific ordinal scaling of brain size. It implies a determination of brain weight by body surfaces rather than by volume, a finding that has not yet been satisfactorily explained[2]. Similarly, it is difficult to explain the geometric similarity between neonatal weight and maternal surface. One explanation may be that, assuming (as for other organs[5]) virtual isometry of placental weight and maternal weight, placental surface rather than volume determines neonatal weight.

Among macaque species, a (0.49 ± 0.03) is clearly below 0.75 and 0.67. As with other aspects of scaling of neonatal size, a parallel situation exists in the scaling of brain size. Studies of brain–body size relationships in primates[10-13] have shown that brain weights in adults of a single species (or a cluster of related species so similar that they appear as different size expressions of the same body plan) scale at a much lower power of the body weight (generally between 0.2 and 0.4) than more distantly related species. As with intrageneric curves of brain size[2], the present findings on macaque species suggest that the low value of a in an intrageneric curve of neonatal size is probably a consequence of the intraspecific pattern which in turn represents the correlated variability on which selection acts.

In summary, the findings on neonatal–maternal weight relationships in eutherian mammals indicate that neonatal size is not scaled metabolically. In taxa that are structurally and functionally equivalent in foetal membranes and placenta, neonatal weight scales at a power of maternal weight close to 0.67. This suggests that maternal (placental) surfaces rather than volume determine neonatal size.

I thank W. P. McNulty, Oregon Regional Primate Research Centre, for weight data for macaque species and S. J. Gould, W. P. Luckett, G. T. Sacher, K. A. Bennett, J. T. Kelly, S. G. Larson and J. A. Vilensky for constructive criticism, and J. Noda for constructing the figure. Support was provided by a grant from the Graduate School, University of Wisconsin, Madison.

WALTER LEUTENEGGER

Department of Anthropology,
University of Wisconsin,
Madison, Wisconsin 53706

1 Leutenegger, W., *Nature*, **240**, 568-569 (1972).
2 Gould, S. J., *Contrib. Primat.*, **5**, 244-292 (1975).
3 Spencer, R. P., *Physiologist*, **14**, 235 (1971).
4 Leutenegger, W., *Folia primat.*, **20**, 280-293 (1973).
5 Stahl, W. R., *Science*, **150**, 1039-1041 (1965).
6 Frick, H., *Z. Säugetierk*, **22**, 193-297 (1957).
7 Frick, H., *Anat. Anz.*, **104**, 17-20, 305-333 (1957).
8 Claussen, C. P., and Schleifenbaum, L., in *Proc. third int. Congr. Primat., Zurich 1970*, **1** (edit. by Biegert, J., and Leutenegger, W.), 122-129 (Karger, Basel, 1971).
9 Luckett, W. P., in *Phylogeny of the Primates: a Multidisciplinary Approach* (edit. by Luckett, W. P., and Szalay, F. S.), 157-182 (Plenum, New York, 1975).
10 Scholl, D., *Proc. R. Soc.*, B135, 243-258 (1948).
11 Bauchot, R., and Stephan, H., *Acta zool.*, **45**, 63-75 (1964).
12 Leutenegger, W., *Folia primat.*, **19**, 9-17 (1973).
13 Pilbeam, D., and Gould, S. J., *Science*, **186**, 892-901 (1974).

Table 1 Scaling of neonatal weight in eutherian mammals and subordinate eutherian taxa

Taxon	Level	N	b	a	μ	r
Eutheria	Infraclass	36	0.34	0.83	—	0.92
Primates	Order	29	0.41	0.78	0.05	0.97
Strepsirhini	Suborder	9	0.64	0.63	0.05	0.99
Haplorhini	Suborder	20	1.08	0.69	0.04	0.99
Macaca	Genus	7	6.22	0.49	0.03	0.97

N, Sample size (number of species); b, initial size constant (y intercept); a, allometric exponent (slope); μ, 95% confidence interval of a; r, correlation coefficient.
—, No value given in literature.

Phenotypic variability of inbred and outbred mice

THE relative phenotypic variability of inbred and commercially available outbred stock mice has never been satisfactorily investigated, even though it has important implications in the design of animal experiments. Early workers[1] suggested that "the published facts suggest that inbred strains are usually more variable than inter-strain hybrids, with random bred colonies occupying an intermediate position." Later it was concluded that with respect to phenotypic variability, "no

general law can be asserted on the comparison of inbred strains with randombred stocks"[2]. A subsequent debate on the suitability of inbred, outbred and F_1 hybrid mice for biological assay[3-6] failed to clarify the situation, and many research workers were left with the impression that inbred animals are more variable than outbred ones. Much of the early work could be criticised on the grounds that sample sizes were small, and few populations were sampled. We have collected extensive data on the phenotypic variability of the shape of the mandible, as part of a routine genetic quality control programme[7,8], and have found that assuming equal sensitivity to an experimental treatment 50% more outbred than inbred mice would be required to detect a 1-unit change in mandible shape.

Samples of mandibles from inbred (I), outbred (O), F_1 (F_1) and F_2 (F_2) hybrid mouse colonies were obtained from a number of locations and colonies, as summarised in Table 1.

The outbred stocks were named as: TO, BKW, CFLP, LACA, LACG, ICI, MF1, CD-1, CFW, Porton and CF_1. They were all commercially available "white mice" (except for the LACG) typical of those used for many research purposes. Inbred strains included the following, A, A2G, AKR, B10.BR, B10.A, B10.LP-*a*, BALB/c, C57BL, CBA/Ca, CBA/H-T6, CFNIH, C3H/He-*mg*, C3H/He, CE, C57BL/10, C57BR/cd, DBA/1, DBA/2, F/St, ICFW, NZW and 129/RrJ. The F_1 hybrid mice were (C3H × DBA/2)F_1 and (BALB/c × C57L)F_1. The F_2 hybrids were (NZB × NZW)F_2 and (BALB/c × C57L)F_2.

Each sample consisted of 8–28 (average 12) male mice over 6 weeks old, and usually weighing 25–30 g, and at least two separate samples were obtained from each colony.

Mandibles were prepared and measured as previously described[9] using 11 measurements per mandible, and the shape of each mandible was described by two discriminant functions

$$d_i = \sum_j \beta_{ij} X_j \qquad \begin{array}{l} i = 1, 2; \\ j = 1, 2 \ldots 11 \end{array}$$

where the β_{ij} are a set of constants given elsewhere (ref. 8, Table 3) calculated from a previous set of data (not included in this study) and the X_j are the individual measurements expressed as a percentage of the sum of the measurements, in order to

correct for variation in size of individuals. The traits d_1 and d_2 may be regarded as numerical indices of mandible shape corrected for size. The data analysed in this study are the within-sample standard deviation of the two discriminant functions d_1 and d_2, designated S_1 and S_2 in the 194 samples from the four different classes of stock. Although d_1 and d_2 discriminate well between strains, they do not discriminate between I, O, F_1 and F_2 classes, and there is no evidence for any association between mean and variance. Thus, the use of these two functions should not introduce any bias when comparing the four classes of stock. Also, d_1 and d_2 are uncorrelated.

Assuming that the two characters d_1 and d_2 have a polygenic mode of inheritance, the variation within the F_1 crosses should give a direct estimate of the environmental variance V_E. The total genetic variance (including additive and non-additive components) within the F_2 and O classes of stock may then be estimated by subtraction.

The distribution of the within-sample standard deviations is shown in Fig. 1 for trait S_1. Within the inbred class of stock there was no evidence that some strains were more variable than others. Similarly, there was no evidence that some of the outbred colonies were more variable than others, though Fig. 1 shows that the variation between samples was greater than with the inbreds, and as some are likely to be more inbred than others, differences could have been expected.

Pooling the data within each class of stock, the following relationships for phenotypic variability were established

$$\text{Trait } S_1: F_2 > {}^*O > {}^{**}I > F_1$$
$$\text{Trait } S_2: F_2 > {}^*O > {}^{**}I > F_1$$

*Significant $(0.01 < P < 0.05)$
**Significant $(P > 0.01)$

The mean phenotypic standard deviation in each class of stock, together with standard errors, is summarised in Table 2. Inbreds were slightly but not significantly more variable than F_1 hybrids, but were substantially less variable than the outbreds, which in turn were less variable than the F_2 hybrids.

Table 1 Animals used in the study

Classification	F_2	O	I	F_1	Total
Locations	1	7	5	5	—
Colonies	2	14	28	2	—
Named stocks	2	11	22	2	37
Samples	10	63	107	14	194
Mice	114	873	1,101	256	2,344

Fig. 1 Distribution of the within-sample phenotypic standard deviation of an index of mandible shape in F_2 hybrid, outbred (O), inbred (I) and F_1 hybrid mice. Arrows indicate the mean, *n* is the number of samples. Sample size averaged 12 mice.

Table 2 Mean phenotypic standard deviation (and its standard error) of two indices of mandible shape ($\bar{S}_1$ and $\bar{S}_2$)

Class of stock	No. of samples	$\bar{S}_1$	s.e.	$\bar{S}_2$	s.e.
F_2	10	1.73	0.161	1.52	0.101
O	63	1.49	0.055	1.39	0.050
I	107	1.14	0.022	1.06	0.030
F_1	14	1.01	0.043	0.97	0.056

Estimates of the environmental variance, and the genetic variance within the F_2 and O stocks is given in Table 3. Heritabilities (in the broad sense) of d_1 and d_2 were estimated as 0.66 ± 0.175 and 0.59 ± 0.216, respectively. As single-cross F_2 stock has a coefficient of inbreeding of 0.5 (ref. 10), the relatively low genetic variance of the O stock compared with the F_2 stock implies a substantial degree of inbreeding, probably equivalent on average, to four or more generations of brother × sister mating.

This study shows unequivocally that for these two highly heritable polygenic characters outbred stocks are substantially more variable than inbreds. The implications in terms of number of animals required in an experiment may easily be

Table 3 Genetic and environmental variances

Variance	Estimated from	Trait S_1	s.e.	Trait S_2	s.e.
V_E	V_{F_1}	1.02	0.061	0.94	0.076
$V_{G_{F_2}}$	$V_{F_2} - V_{F_1}$	1.97	0.398	1.37	0.230
$V_{G_0^{F_2}}$	$V_0 - V_{F_1}$	1.20	0.129	0.99	0.124

calculated on the assumption that none of the classes of stock is likely to be more sensitive to an experimental treatment than another. In these circumstances, an experiment set-up to detect a 1 unit change in mandible shape between treated and control groups (assuming type 1 and 2 errors are set at 5%, and 10% respectively) would require 13 F_1, 16 I, 25 O or 34 F_2 animals. Thus, in this case, more than 50% more outbred than inbred animals would be required, though the difference would not be so large in the case of characters with a lower heritability.

These findings give added support to the case for the more widespread use of inbred strains[11].

I thank Mr W. Peacock for technical assistance.

Michael F. W. Festing

Medical Research Council Laboratory
Animals Centre,
Woodmansterne Road,
Carshalton, Surrey SM5 4EF, UK

1 McLaren, A., and Michie, D., *J. Genet.*, **54**, 440–445 (1956).
2 Biggers, J. D., McLaren, A., and Michie, D., *Nature*, **182**, 77–90 (1958).
3 Chai, C. K., *Nature*, **185**, 514–518 (1960).
4 Biggers, J. D., McLaren, A., and Michie, D., *Nature*, **190**, 891–892 (1961).
5 McLaren, A., and Michie, D., *Nature*, **190**, 894 (1961).
6 Becker, W. A., *Nature*, **193**, 1264–1266 (1962).
7 Festing, M. F. W., *Lab. Anim.*, **8**, 265–270 (1974).
8 Festing, M. F. W., *Lab. Anim.*, **8**, 291–299 (1974).
9 Festing, M. F. W., *Nature*, **238**, 351–352 (1972).
10 Falconer, D. S., *Introduction to Quantitative Genetics*, chs. 8 and 15, and p. 95 (Oliver and Boyd, Edinburgh, 1960).
11 Festing, M. F. W., *Food and Cosmetics Toxicology*, **13**, 369–375 (1975).

Fine structural changes in human astrocyte carrier lines for measles virus

Chronic infection with measles virus is causally related to the human syndrome of subacute sclerosing panencephalitis (SSPE)[1-5], and means of reproducing that condition in experimental animals are being sought[6-10]. Measles carrier states *in vitro* have been established in some human[11-13] and rodent[14,15] cell systems. We now report the establishment of measles carrier states in human astrocyte cultures of normal brain, and in a tumour astrocyte line, as well as the emergence of fine structural features which bear a close resemblance to findings in the brain of some patients with SSPE[16].

Two strains of measles virus were used—wild type (WT) derived from a pharyngeal swabbing of a patient with measles, and Edmonston (ED), a laboratory strain[17]. Infectivity (TCID$_{50}$/0.1 ml of cell-free supernatant), haemagglutination (HA) and neutralisation titrations followed published procedures[17-19], as did primary and secondary cell culture techniques[20]. The same batches of Eagle's MEM and of foetal calf serum were used throughout.

In initial work to establish measles carriers we used the 1N1 tumour astrocyte line[21]; equivalent 1N1 cells had been used successfully to initiate the V-1N1 line of visna virus-producing human astrocytes[22]. Subsequent work with astrocytes derived from two normal foetal brains (HFA1, HFA2) and with WT virus gave an essentially similar pattern of virus–cell interaction. The following description of the establishment of the ED-human tumour astrocyte line (ED-1N1) will therefore serve, save for minor variations of time scale, for the ED carriers derived from normal astro-cytes (ED-HFA1; ED-HFA2) and the WT astrocyte carriers (WT-1N1; WT-HFA1; WT-HFA2).

1N1 cells at passage 125 were used. At this passage level, the cells contain glycogen[21], synthesise high levels of adenyl cyclase[23] and respond appropriately to catechol stimulation[24], as befits astrocytes. 1N1 cultures at a density of 1×10^4 cells per mm² were exposed at 37 °C to 0.01 TCID$_{50}$ ED virus per cell for 2 h, then rinsed and returned to growth medium. Binuclear cells were plentiful in 2 d, small syncytia in 3, huge syncytia in 5; cytolysis began on about day 5 and was almost total by day 9. Intranuclear and cytoplasmic inclusions were made visible by Giemsa staining. Infectivity titres at day 7 after infection were $10^{4.5}$ TCID$_{50}$ and HA was 14–28. These results are in accord with an acute cytolytic infection of classical type by measles virus[17,19].

In occasional cultures, scattered cells survived for 14 d; such survivors were likely to develop into carriers. The ED-1N1 line was derived from cells infected as above: a few discrete cells survived for 9 d after infection; these were elongated and had prominent paranuclear vacuoles. At 14 d after infection five small cell groups were present; 25 d after infection mitoses were frequent; the most common cell was plump and fusiform. Small syncytia appeared 31 d after infection among the other cells, grew slowly to maximum size (always much smaller than in cytolytic states) in 8 d and eventually rolled off, while other syncytia developed. The predominant cell was mononuclear. At 60 d after infection the cells were trypsinised for the first time, and designated

Fig. 1 *a*, Intranuclear inclusion, as seen in cytolytic cultures of 1N1 cells 5 d after infection with ED measles virus. Higher magnifications (*b*) show them to consist of arrays of 16-nm smooth nucleocapsid tubules. From fourth transfer onwards, such tubules were not observed in the carrier lines. The bar represents 500 nm.
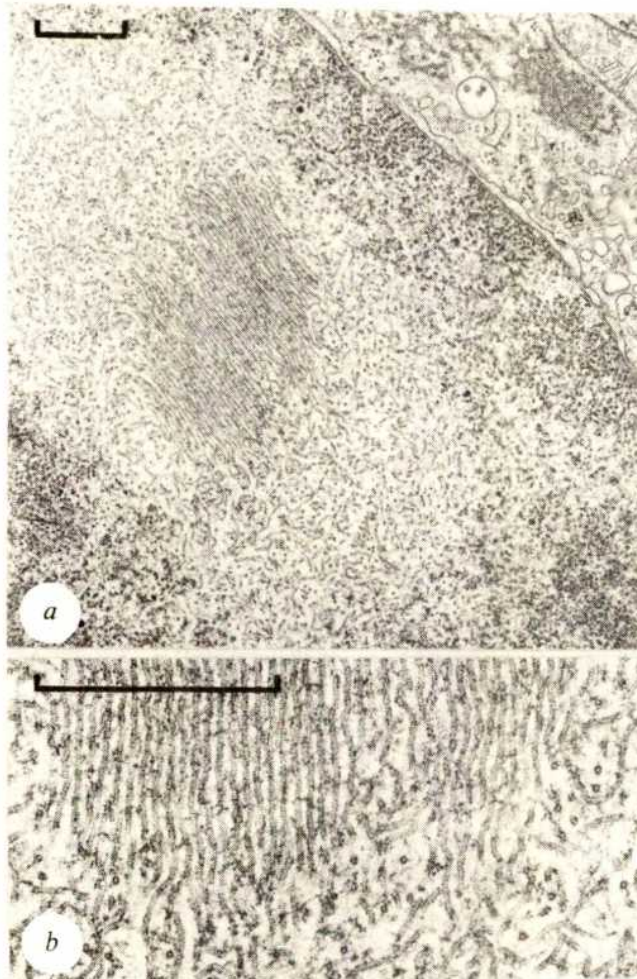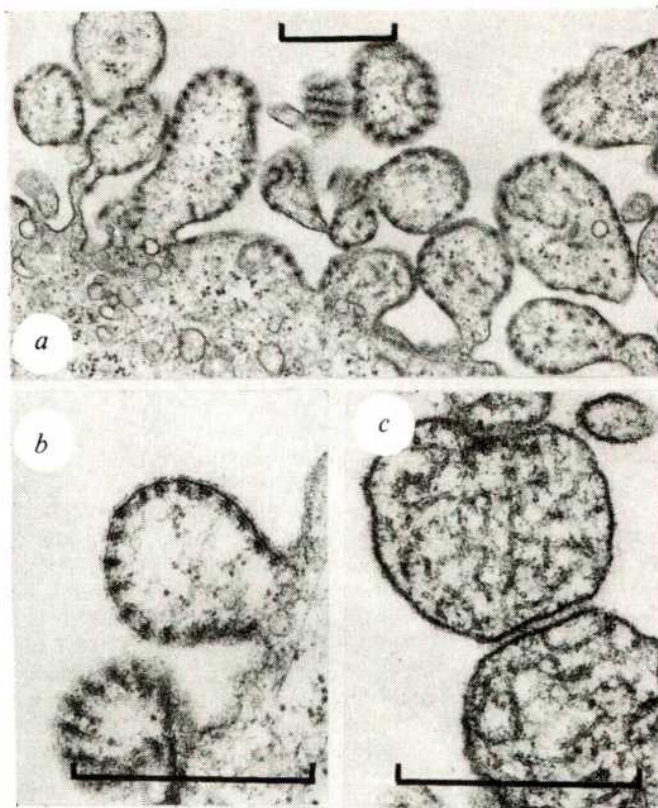
Fig. 2 *a* and *b*, Development and structural detail of complete measles virions at the cell surface, as found in both cytolytic and carrier cultures. *c*, Structurally defective particles observed (beside complete virions) from fifteenth transfer stage in the carrier lines. Note the lack of external projections and disorderly packaging of the nucleocapsids. The bar represents 500 nm.

ED-1N1[?]. Intranuclear and cytoplasmic inclusions were present in both fusiform cells and syncytia. Infectivity titres at this stage were 10^3 TCID$_{50}$. Between transfers 5 and 15 they were consistently 10^5 TCID$_{50}$, while HA titres were 28–56 and anti-measles neutralisation stood at 1/4. In later transfers, interference was evident, that is, no syncytia were expressed at 10^0, 10^{-1} and $10^{-1.5}$, but they were seen at high dilutions of virus. Morphologically atypical virus was detected subsequently in electron micrographs of such cultures. So far the ED-1N1 line has had 21 trypsinisations; this covers 32 weeks from initial trypsinisation. The cells continue to grow well, to produce inclusions and scattered syncytia, and virus titres now stand at $10^{3.5}$ TCID$_{50}$. The HA titres of the WT lines (WT-1N1; WT-HFA1; WT-HFA2) are 128–256; those of the ED lines (ED-1N1; ED-HEA1; ED-HEA2) are 14–28. The WT lines are at transfer 12.

For electron microscopy[21,25] representative cultures of the 1N1 cells 5 d after infection with ED strain and from transfers 3 to 21 of the noncytolytic ED-1N1 carrier line were compared with equivalent transfers of uninfected HFA1, HFA2 and 1N1 cells, and with WT-1N1[6,12], WT-HFA1[6], WT-HFA2[6] and ED-HFA1[6] carrier lines.

The ultrastructure of the cytolytic cultures showed all the usual intranuclear and cytoplasmic features (see for example, refs 26 and 27) (Fig. 1*a* and *b*). Extracellular virions featured a surface layer of 8-nm projections, characteristically thickened envelope membrane, and nucleocapsid tubules disposed peripherally in spiral form (Fig. 2*a* and *b*). The infected nuclei also had many rounded condensations of microfilamentous material resembling the "nuclear bodies" of normal cells[28].

Between transfers 3 and 21 of the ED-1N1 carrier line there were three morphological changes: (1) the dis-

appearance of measles intranuclear nucleocapsids, (2) the appearance in the nuclei of complex structures seemingly derived from nuclear bodies and (3) the emergence of a proportion of morphologically defective virus particles at the cell surfaces. Intranuclear aggregates of measles nucleocapsid[26,27] were not found in any of the carrier cultures. In rare cells at transfer three scattered tubular profiles were found whose size matched that of smooth nucleocapsid. From the fourth transfer onwards, even these occasional structures were absent. At this stage, however, unusually large nuclear bodies (up to 800 nm diameter) became frequent, some with centrally placed clusters of vesicles (Fig. 3*b*). By transfer 7, enlarged, structurally abnormal nuclear bodies, from 500 nm to 3.5 μm in diameter and with a well developed membranous component, were a feature of almost all the cells. Four types could be recognised, perhaps representing different developmental stages: (1) simple microfilamentous, (2) vesicular, (3) complex laminar and (4) cisternal (Fig. 3*a–d*). A single nucleus often held different types. Such bodies were invariably demarcated from surrounding nuclear structures by a peripheral zone of microfilamentous matrix, like that of nuclear bodies in control cells[28]. The new structures were easily distinguished from nucleoli or infoldings from the nuclear envelope. Most striking were the large laminar and cisternal types, in which paired membranes formed closely coiled profiles, with various degrees of focal distension or ballooning. A layer of fine (9–10 nm) regularly spaced filaments often was intercalated between the paired membranes (Fig. 3*c*). Such modified nuclear bodies, and an absence of measles intranuclear nucleocapsid tubules, also characterised the other carrier lines. Extensive surveys of uninfected cultures showed normal small nuclear bodies; no suggestion of the more complex structures was ever seen.

All carrier lines had numerous, large cytoplasmic inclusions akin to those found in cytolytic cultures. Virus particle morphogenesis was evident in all carriers; but in later transfer stages, especially the 15th and 21st transfers

Fig. 3 Four types of nuclear bodies, present within nuclei of ED-1N1 carrier cultures at fourth and seventh transfer stages: *a*, simple microfilamentous; *b*, vesicular; *c*, complex laminar; *d*, cisternal. The more elaborate forms were a prominent feature from the seventh transfer stage onwards. The bar represents 500 nm.
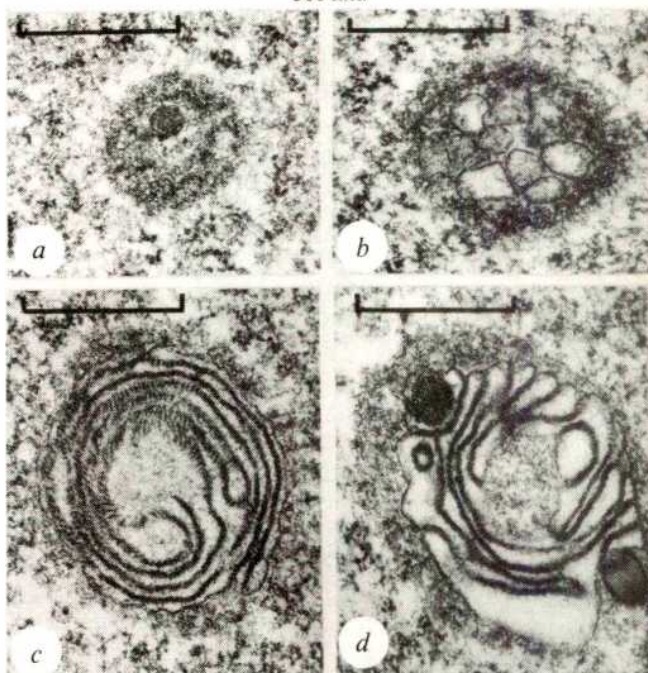
of the ED-1N1 line, other types of membrane-bound particles were also found (Fig. 2c); their envelopes resembled unmodified host plasma membrane, lacking external projections or thickening of the unit membrane. The contained nucleocapsid tubules lacked the characteristic peripheral and spiral disposition of complete measles virions (Fig. 2a and b). These were considered to be morphologically defective measles particles: their presence would seem to complement the observation of viral interference noted in infectivity titrations.

We conclude that the measles carrier lines exhibited a common pattern of fine structural changes, whether derived from normal or tumour astrocytes and whether infected with ED or WT virus. It would be interesting to know whether measles vaccine preparations can induce a similar *in vitro* response in suitable host cells. A prominent feature of the carrier lines, not reported previously even in long term measles-infected organ cultures[14], was the emergence of complex intranuclear formations interpreted as modified nuclear bodies. Martinez *et al.*[16], in a review based on eight autopsied cases of SSPE, defined a category of nuclear inclusions believed to be of diagnostic significance and representing altered nuclear bodies, distinct from paramyxovirus-like nucleocapsid tubules also reported in this disease[1-5,26,27]. Enlarged nuclear bodies of the same kind were in evidence from the third transfer onwards in our astrocyte carrier lines, together with even more elaborate forms with a well developed membranous component. Simple hypertrophy of nuclear bodies is well known in measles, and some other virus infections[28], but their relationship, if any, to synthesis of specific viral components is still largely speculative. SSPE is associated with a defective form of measles virus[4]. Although the carrier lines continued to produce infective virus throughout, evidence of defective particle production was also obtained.

Functional and structural features of the measles carrier state in human astrocytes seem therefore to present interesting parallels with the disease SSPE.

E. H. Macintyre
J. A. Armstrong

National Institute for Medical Research,
Mill Hill,
London, NW7 1AA, UK

1 Bouteille, M., Fontaine, C., Vedrenne, C. L., and Delarue, J., *Rev. Neurol.*, 118, 454–458 (1965).
2 Horta-Barbosa, L., Fuccillo, D. A., Sever, J. L., and Zeman, W., *Nature*, 221, 974 (1969).
3 Connolly, J. H., Haire, M., and Hadden, D. S. M., *Br. med. J.*, 1, 23–25 (1971).
4 ter Meulen, V., Katz, M., and Müller, D., *Curr. Top. Microbiol. Immun.*, 57, 1–38 (1972).
5 Dubois-Dalcq, M., Barbosa, L. H., Hamilton, R., and Sever, J. L., *Lab. Invest.*, 30, 241–250 (1974).
6 Wear, D. J., Rabin, E. R., Richardson, L. S., and Rapp, F., *Expl molec. Path.*, 9, 405–417 (1968).
7 Baringer, J. R., and Griffith, J. F., *Lab. Invest.*, 23, 335–346 (1970).
8 Wear, D. J., and Rapp, F., *J. Immun.*, 107, 1593–1598 (1971).
9 Thormar, H., Jervis, G. A., Karl, S. C., and Brown, H., *J. infect. Dis.*, 127, 678–685 (1973).
10 Raine, C. S., Bynington, D. P., and Johnson, K. P., *Lab. Invest.*, 31, 355–368 (1974).
11 Rustigian, R., *Virology*, 16, 101–104 (1962).
12 Norrby, E., *Arch. Virusforsch.*, 20, 215–224 (1967).
13 Webb, H. E., Illavia, S. J., and Laurence, G. D., *Lancet*, ii, 4–5 (1971).
14 Raine, C. S., Feldman, L. A., Sheppard, R. D., and Bornstein, M. B., *J. Virol.*, 8, 318–329 (1971).
15 Knight, P., Duff, R., and Rapp, F., *J. Virol.*, 10, 995–1001 (1972).
16 Martinez, A. J., Ohya, T., Jabbour, J. T., and Dueñas, D., *Acta Neuropath., Berl.*, 28, 1–13 (1974).
17 Enders, J. F., and Peebles, T. C., *Proc. Soc. exp. Biol. Med.*, 86, 277–286 (1954).
18 Norrby, E., *Proc. Soc. exp. Biol. Med.*, 111, 814–818 (1962).
19 Katz, S. L., and Enders, J. F., in *Diagnostic Procedures for Viral and Rickettsial Infections*, fourth ed. (edit. by Lennette, E. H., and Schmidt, N. J.), 504–528 (American Public Health Association, New York, 1969).
20 Pontén, J., and Macintyre, E. H., *Acta path. microbiol. scand.*, 74, 465–486 (1968).
21 Macintyre, E. H., Pontén, J., and Vatter, A. E., *Acta path. microbiol. scand.*, 80A, 267–283 (1972).
22 Macintyre, E. H., Wintersgill, C. J., and Thormar, H., *Nature new Biol.*, 237, 111–113 (1972).
23 Perkins, J. P., Macintyre, E. H., Riley, W. D., and Clark, R. B., *Life Sci.*, 10, 1069–1080 (1971).
24 Clark, R. B., and Perkins, J. P., *Proc. natn. Acad. Sci. U.S.A.*, 68, 2757–2760 (1971).
25 Hirsch, J. G., and Fedorko, M. E., *J. Cell Biol.*, 38, 615–627 (1968).
26 Nakai, T., Shand, F. L., and Howatson, A. F., *Virology*, 38, 50–67 (1969).
27 Raine, C. S., Feldman, L. A., Sheppard, R. D., and Bornstein, M. B., *J. Virol.*, 4, 169–181 (1969).
28 Bouteille, M., Kalifat, S. R., and Delarue, J., *J. Ultrastruct. Res.*, 19, 474–486 (1967).

Environmental carcinogens and large bowel cancer

The geographical variation in the incidence of large bowel cancer together with the increased incidence found in migrants from low risk to high risk areas support the view that environmental factors are involved in the aetiology of this disease[1]. These environmental factors are thought to be largely dietary although other influences have not been excluded. Aries *et al.*[2] suggested that the disease is caused by carcinogens produced by gut bacteria from bile acids. Diet would influence both the bile acid levels in the bowel and also the nature of the gut flora. Faecal bile acids are more concentrated in faeces from people living in high incidence areas than in faeces from populations at low risk[3]. Similarly, faecal bile acids are usually more concentrated in samples from patients with large bowel cancer than in those from control patients[4]. Bile acids were postulated to be the substrates from which the gut flora produced carcinogens[2] and indeed this suggestion has been supported by the demonstration that the faecal concentration of bacterially degraded bile salts is more highly correlated with the incidence of large bowel cancer than is the total bile acid level[3]. However, it has been suggested that the primary role of bile acids is as co-carcinogens and this suggestion is supported by the results of animal experimentation demonstrating the promoting effect of bile acids on the action of known carcinogens[5]. If it is accepted that bile acids act as co-carcinogens it becomes necessary to explain how the carcinogen reaches the large bowel without producing cancers in other parts of the gastrointestinal tract. Here we describe preliminary investigations of a mechanism by which an environmental carcinogen could be transported to the large intestine.

The conjugates and metabolites of benzo[a]pyrene were prepared biologically. This compound has been shown to undergo oxidation and conjugation in the rat and is excreted mainly in the bile as glucuronic acid and sulphate conjugates of phenols, as conjugates of dihydrodiols and possibly as mercapturic acid conjugates[6,7]. Two bile duct cannulated female Wistar albino rats were given 7,10,14-^{14}C-benzo[a]pyrene (Radiochemical Centre, Amersham); 1 mg; 14 μCi per rat. The bile was collected for 4 h and found to contain 6.7 and 12.5% of the administered ^{14}C. Analysis by thin layer chromatography[8] showed a single peak of radioactivity at the origin, corresponding to conjugated metabolites. Aliquots of the pooled bile containing the ^{14}C metabolites of benzo[a]pyrene were incubated with samples of human and rat faeces homogenised in nutrient broth. Further aliquots of the bile were incubated with pure cultures of bacteria representing the major types found in the intestine. After incubation for 72–96 h under strictly anaerobic conditions the reaction mixtures were analysed by thin layer chromatography.

Almost all the bacteria tested were able to hydrolyse the conjugates to release the oxidative metabolites of benzo[a]pyrene (Table 1) and the major conjugated metabolites of benzo[a]pyrene seem to be phenols and quinones. It is, however, possible that further metabolism of the unconjugated metabolites may have occurred, and these phenols may have been produced from or converted to dihydrodoils and, to some extent, to benzo[a]pyrene. Of particular interest are the findings that the organism releasing the greatest amount of the parent hydrocarbon in this test system, *Clostridium paraputrificum*, is present in larger numbers in the faeces of "western" subjects[9] than in faeces from low risk areas, and the *Bacteroides* spp. show this activity as well as deconjugation of taurocholate and dehydroxylation of bile acids.

Table 1 The bacterial hydrolysis of the conjugates of benzo[a]pyrene excreted in the bile of rats

| | | Proportion of benzo[a]pyrene and its metabolites | | |
	Conjugates hydrolysed (%)	Dihydrodiols	Phenols and quinones	Benzo[a]pyrene
Rat faeces	28.0	15.9	11.3	0.7
Human faeces	28.4	10.2	17.1	0.9
B. fragilis 9343	14.6	2.2	11.8	0.8
B. fragilis ST 2	26.0	7.2	18.4	0.4
Cl. perfringens 10379	12.6	1.9	10.4	0.3
St. faecalis 775	13.5	2.6	10.5	0.4
E. coli	10.0	3.1	6.8	0.1
Kl. pneumonea	13.2	2.0	10.5	0.7
Cl. sporogenes	8.9	0	8.9	0.3
Cl. sporogenes	0	0	0	0.3
Cl. paraputrificum	12.3	0.1	11.4	0.9

The results are expressed as the percentage of the total radioactivity in the incubate.

Rat bile (100 μl) containing the conjugated metabolites of benzo[a]pyrene was incubated at 37 °C under anaerobic conditions for 72–96 h. Incubation in the absence of microorganisms showed a limited hydrolysis and conversion to benzo[a]pyrene, that is 10.0 and 0.5% respectively, which has been subtracted from the bacterial incubations to give the above results.

These preliminary experiments show that bacteria of the types found in the human intestine are able to hydrolyse the conjugated metabolites of benzo[a]pyrene excreted in the bile of the rat. A small, but significant, amount of benzo[a]pyrene was also released, probably as a result of the dehydroxylation of hydroxylated metabolites. These reactions are a reversal of the detoxication processes performed by the liver. An analogous mechanism is thought to explain the intestinal cancers produced in rats as a result of the administration of 3-methyl-4-aminobiphenyl and 3-methyl-2-aminonaphthalene, and the relevance of such studies with aromatic amines to large bowel cancer was reviewed by Weisburger[10].

Polycyclic aromatic hydrocarbons, such as benzo[a]pyrene, are widely distributed in nature and arise mainly from the combustion of organic matter, especially fossil fuels. Their concentration in the environment is thus increased by industrial activity. Engst and Fritz[11] showed concentrations of 0.8 μg kg^{-1} in sea sand, up to 0.8 mg kg^{-1} in rural soils but 65 mg kg^{-1} in the industrial urban environment. These compounds also occur in smoked foods. We suggest that polycyclic aromatic hydrocarbons such as benzo[a]pyrene ingested from the environment may act as the primary carcinogens in the development of large bowel cancer. The geographical variations in the incidence of this disease could then be partly explained in terms of the variations in the distribution of the carcinogens associated with industrial activity, since the incidence of colon and rectal cancer is highly correlated with gross national product[12].

The consumption of a western diet, rich in fat and protein, is undoubtedly also important. Diet probably influences the metabolism of polycyclic aromatic hydrocarbons directly, as a result of its influence on mucosal enzymes and indirectly by the effect on bile acid metabolism. Aryl hydrocarbon hydroxylase activity is present in the duodenal mucosa of rats fed fat-containing diets and may be induced by exposure to benzanthracene. The enzyme is present at very low levels in the mucosa of rats fed fat-free diets and is detected in the colon of rats fed fat-containing diets only after exposure to benzanthracene[13]. The levels of other enzymes, associated with benzo[a]pyrene metabolism in the gut wall, have not been extensively investigated.

There are two ways in which bile acids may be important in the metabolism of polycyclic aromatic carcinogens excreted in the bile. (1) The conjugated metabolites of benzo[a]pyrene in the bile are amphipathic compounds, as are the bile acids. An important function of bile acids is the formation of mixed micelles during fat digestion and it is possible that polar conjugates are incorporated into micelles thus aiding their transport through the small intestine. (2) Polycyclic aromatic hydrocarbons are lipid soluble and it has been suggested that this property would render them inactive in the aqueous environment of the gut[14]. Bile acids are,

however, powerful solubilising agents and would maintain the carcinogens in solution and aid their interaction with cell surfaces.

In conclusion, we would suggest that large bowel cancer results from the retoxification by the gut flora of the liver metabolites of polycyclic aromatic hydrocarbons. This process is influenced by the effect of diet on the intestinal cells and secretions, particularly bile acids and possibly by the effect of diet on the gut flora. The disease is more common in western countries because (1) industrial activity increases the amounts of polycyclic aromatic carcinogens in the environment and (2) the western diet favours the retoxification process.

A. G. RENWICK

Faculty of Medicine, University of Southampton,
Medical and Biological Sciences Building,
Bassett Crescent East, Southampton SO9 3TU, UK

B. S. DRASAR

Bacterial Metabolism Research Laboratory,
Rear of Colindale Hospital,
Colindale Avenue, London NW9 5DX, UK

1 McIllmurray, M. B., and Langman, M. J. S., *Gut*, 17, 815–820 (1975).
2 Aries, V. C., Crowther, J. S., Drasar, B. S., Hill, M. J., and Williams, R. E. O., *Gut*, 10, 334–335 (1969).
3 Hill, M. J., Drasar, B. S., Aries, V. C., Crowther, J. S., Hawksworth, G., and Williams, R. E. O., *Lancet*, i, 95–100 (1971).
4 Hill, M. J., Drasar, B. S., Williams, R. E. O., Meade, T. W., Cox, A., Simpson, G., and Morson, B. C., *Lancet*, i, 535–538 (1975).
5 Reddy, B. A., Narisawa, T., Maronpot, R., Weisburger, J. H., and Wynder, E. L., *Cancer Res.*, 35, 3421–3426 (1975).
6 Kotin, P., Falk, H. L., and Busser, R., *J. natn. Cancer Inst.*, 23, 541–555 (1959).
7 Falk, H. L., Kotin, P., Lee, S. S., and Nathan, A., *J. natn. Cancer Inst.*, 28, 699–724 (1962).
8 Sims, P., *Biochem. Pharmac.*, 16, 613–618 (1967).
9 Drasar, B. S., Goddard, P., Heaton, S., Peach, S., and West, B., *J. med. Microbiol.*, 9, 63–71 (1976).
10 Weisburger, J. H., *Cancer*, 28, 60–70 (1971).
11 Engst, R., and Fritz, W., *Proc. 9th Congr. Nutrition*, 3, 303 (Karger, Basel, 1975).
12 Armstrong, B., and Doll, R., *Int. J. Cancer*, 15, 617–631 (1975).
13 Wattenberg, L. W., Leong, J. L., and Strand, P. J., *Cancer Res.*, 22, 1120–1125 (1962).
14 Wynder, E. L., and Shigematsu, T., *Cancer*, 20, 1520–1561 (1967).

T–T cell collaboration during *in vivo* responses to antigens coded by the peripheral and central region of the MHC

MIXED lymphocyte culture (MLC)[1] has been used extensively as an *in vitro* model to analyse the reactivity of T cells to antigens coded by the major histocompatibility complex (MHC). When murine T responder cells are exposed *in vitro* to allogeneic lymphoid cells (stimulator cells) they proliferate and cytotoxic T lymphocytes (CTL) are generated[2,3]. Antigens coded by the central I region of the MHC are chiefly responsible for triggering prolifer-

Table 1 Collaborative effect of *in vitro* responses to antigens coded by the peripheral and central region of the H-2 complex

Responder cell	Stimulator cell (X irradiated)	H-2 incompatibility K,IA,IB,IC,D	^{3}H-thymidine uptake (c.p.m.)	% Specific lysis of targets CBA (k kkk k)	ATH (s sss d)
	ATL (s kkk d)	— — — — —	400 ±80	-1	-3
	ATH (s sss d)	— + + + —	22,000 ±3,600	6	24
ATL (s kkk d)	CBA (k kkk k)	+ — — — +	8,000 ±1,700	32	5
	ATH (s sss d) plus CBA (k kkk k)	— + + + — plus + — — — +	27,000 ±3,100	87	35
	(CBA×ATH)F$_1$ $\left(\dfrac{s\ \ s\ \ s\ \ d}{k\ \ k\ k\ \ k}\right)$	+ + + + +	29,000 ±2,600	92	34

The MLC system used to assay for proliferative responses (^{3}H-thymidine uptake technique) is that described by Peck *et al.*[21]. Splenic responder cells (0.25×10^6 cells per well) were incubated with 0.25×10^6 X-irradiated (2,000 rad) splenic stimulator cells in MLC medium[21] supplemented with 0.5% fresh mouse serum and 5×10^{-5} M 2-mercaptoethanol (ME) in flat bottom microtitre plates (Greiner, Nurtingen/Germany) over a period of 3.5 d. The cells were then pulsed with 1 μCi ^{3}H-thymidine per well for 7 h and processed as described[22]. Six replicate cultures per group were formed. The data given represent the mean ± s.d. The culture system used for the development of CTL has been described elsewhere[7,9]. Briefly, 3×10^6 responder cells were cultured together with 1.5×10^6 irradiated stimulator cells in Marbrook flasks containing MLC medium supplemented with 5% FCS and 5×10^{-5} M ME. After 5 d the cultured cells were titrated for cytotoxicity against LPS-induced blast lymphocytes[9] in a ^{51}Cr assay as described[7,9]. The lysis values given (mean of three determinations) were obtained at a ratio of attacker cells to target cells of 30 to 1: the data are given without s.d., since in the experiments listed the s.d. of percentage specific lysis was less than |4| %. Background lysis during the 3-h assay was less than 18%. The letters within the brackets denote the H-2 genotype of the respective cells.

ation[4,5], whereas the target antigen of the CTL generated is either a H-2K or H-2D region or a I–A subregion gene product[5-8]. This dichotomy in the antigenic requirement of a MLC seems to be reflected at the level of the responding T lymphocytes. Two distinct subsets of T cells (T_1/T_2), which have been defined by both functional[9,10] and physical[11] criteria and possibly on the basis of their Ly phenotype[12], seem to act synergistically during the *in vitro* generation of CTL[9-13]. Proliferating T_1 helper cells are thought to respond mainly to I-region gene products[2,5,12] whereas the precursors of CTL (T_2 cells) are believed to react specifically to transplantation antigens coded by the H-2K, H-2D, or I–A region[2,5,7]. The proliferating T_1 helper cell is believed to potentiate the development of CTL either by itself or through a secreted helper factor[2,11,13,14]. A serious limitation of this concept is, however, that so far the experimental evidence available is derived from work performed *in vitro*. We therefore aimed at verifying T–T collaboration during *in vivo* responses to antigens of the MHC.

The mouse strains we used were chiefly strains ATL, ATH and CBA/J. CBA/J and ATL mice are identical for the H-2I region, whereas ATL and ATH mice share the same H-2K and H-2D regions, but differ at the I region[15]. The results presented in Table 1 were obtained *in vitro* and demonstrate the effect of allogeneic, I-region-coded determinants on the generation of CTL specific for antigens coded by the H-2K and H-2D region. Thus ATL responder T cells proliferate strongly in response to I-region-incompatible ATH stimulator cells. In this combination no CTL reactive against antigens of the H-2K or H-2D region of the H-2^k haplotype are generated (the CTL generated are reactive against antigens of the I–A^k subregion[7]). On the other hand, ATL responder cells respond with a moderate cell proliferation to H-2K and H-2D region-incompatible CBA stimulator cells, and anti-H-2K^k and anti-H-2D^k region-specific CTL are generated. In the presence of "third party" ATH stimulator cells, however, the generation of H-2K^k and H-2D^k region-specific CTL is strongly enhanced. These results basically confirm the work of Bach and associates[5,13,16] and stress the collaborative effect of I-region-coded determinants on the *in vitro* generation of anti-H-2K and anti-H-2D region-specific CTL.

To find out whether during the induction of an *in vivo* allograft reaction a similar collaborative mechanism is operating, the following experimental protocol was applied:

a fixed number of distinct H-2 region-incompatible stimulator cells were injected into the hind footpad of ATL mice. After 3 d the draining lymph nodes were isolated, teased and single-cell suspensions were prepared. The cells were cultured (without adding stimulator cells) for 48–72 h more at 37 °C, a period required to allow full differentiation of the sensitised CTL precursor (A. St-P., in preparation). After this the lymph node cells were tested for the presence of CTL in a ^{51}Cr-cytotoxicity assay. Results of a typical experiment are given in Table 2 and indicate *in vivo* an even more pronounced amplification effect of I-region incompatibility on the generation of H-2K and H-2D specific CTL compared with the *in vitro* data (Table 1). Thus the capacity of H-2K and H-2D region-incompatible stimulator cells to provoke a specific CTL response was greatly increased

Table 2 Collaborative effect of *in vivo* responses to antigens coded by the peripheral and central region of the H-2 complex

Recipient	Irradiated stimulator cells (10^7)	H-2 incompatibility K,IA,IB,IC,D	% Specific lysis of targets CBA (k kkk k)	ATH (s sss d)
	ATL (s kkk d)	— — — — —	0	1
ATL (s kkk d)	ATH (s sss d)	— + + + —	4	19
	CBA (k kkk k)	+ — — — +	12	3
	ATH (s sss d) plus CBA (k kkk k)	— + + + — plus + — — — +	76	27
	(CBA×ATH)F$_1$ $\left(\dfrac{s\ \ s\ \ s\ \ d}{k\ \ k\ k\ \ k}\right)$	+ + + + +	69	24

ATL mice (recipients) were injected with either 10^7 ATL, or ATH, or CBA, or F$_1$ (CBA×ATH), or a mixture of both 10^7 CBA and 10^7 ATH mouse derived irradiated splenic lymphocytes into the left hind footpad. After 3 d the draining lymph nodes (LN) were dissected, single-cell suspensions were prepared and the cells were cultured for 3 more days *in vitro*. In general about 4×10^6–6×10^6 cells were obtained per LN. Thereafter, the cells were titrated for cytotoxicity in a ^{51}Cr assay[7] against LPS-induced target cells[7]. The percentage specific lysis values given were obtained at a ratio of attacker cells to target cells of 30 to 1. Collaborative effects of I region-incompatible stimulator cells were obtained with as few as 2×10^6 ATH stimulator cells injected together with 10^7 CBA stimulator cells (unpublished results). Background lysis of the target cells was less than 16%.

provided I region-incompatible stimulator cells were simultaneously injected into the hind footpad. In other words, the magnitude of the CTL response generated towards antigens coded by the peripheral K or D region of the H-2 complex was amplified by an I region-dependent stimulus presented separately on a "third party" cell.

From the experiments, we conclude that the collaborative effect of a T-cell response to I region-incompatible allogeneic stimulator cells on the generation of anti-H-2K and anti-H-2D-specific CTL seems to take place *in vivo* as effectively as *in vitro*. Since the target antigens of CTL are believed to be the transplantation antigens coded by the H-2K and H-2D region and I-A subregion of the H-2 complex[5,7], and since CTL are thought to be primarily responsible for allograft rejection *in vivo*[3], the *in vivo* immunogenicity of grafted allogeneic cells seems to be strongly dependent on whether or not they express the phenotypical products coded by the I region of the H-2 complex, besides H-2 transplantation antigens. Because lymphoid cells express antigens coded by both the peripheral and central region of the H-2 complex, our results are compatible with the concept that passenger leukocytes[17] may heighten the immunogenicity of transplanted allogeneic tissue grafts. They also offer an explanation of why anti-Ia antibodies eventually result in graft prolongation[18], and are compatible with the finding that allogeneic thyroid transplants are rapidly rejected, provided the recipient's T lymphocytes are sensitised against I region coded cell surface antigens[19,20].

H. WAGNER
A. STARZINSKI-POWITZ
K. PFIZENMAIER
M. ROLLINGHOFF

Institut für Med. Mikrobiologie de Johannes-Gutenberg-Universität,
Hochhaus am Augustusplatz,
D-6500 Mainz, Germany

1 Häyry, P., and Defendi, V., *Science*, **168**, 133–134 (1970).
2 Wagner, H., Röllinghoff, M., and Nossal, G. J. V., *Transplant. Rev.*, **17**, 3–36 (1973).
3 Cerottini, J., and Brunner, K. T., *Adv. Immun.*, **18**, 67–132 (1974).
4 Alter, B. J., *et al.*, *J. exp. Med.*, **137**, 1303–1309 (1973).
5 Schendel, D. J., and Bach, F. H., *J. exp. Med.*, **140**, 1534–1547 (1974).
6 Nabholz, M., Vives, J., Young, H. M., Miggiano, M. V., Rijnbeck, A., and Shreffler, D. C., *Eur. J. Immun.*, **4**, 378–387 (1974).
7 Wagner, H., Goetze, D., Pschelinzew, L., and Röllinghoff, M., *J. exp. Med.*, **142**, 1477–1488 (1975).
8 Hauptfield, V., Klein, D., and Klein, J., *Science*, **181**, 167–169 (1973).
9 Wagner, H., *J. exp. Med.*, **138**, 1379–1397 (1973).
10 Cantor, H., and Simpson, L., *Eur. J. Immun.*, **5**, 620–624 (1975).
11 Wagner, H., Röllinghoff, M., and Shortman, K., in *Progress in Immunology II*, **3**, (edit. by Brent, L., and Holborow, J.), 111–120 (North-Holland Publishing Co., Amsterdam, 1974).
12 Cantor, H., and Boyse, E. A., *J. exp. Med.*, **141**, 1390–1399 (1975).
13 Bach, F. H., Segall, M., Zier, K. S., Sudel, P. M., Alter, F. J., and Bach, M. L., *Science*, **180**, 403–407 (1973).
14 Altmann, A., and Cohen, I. R., *Eur. J. Immun.*, **5**, 437–440 (1975).
15 Shreffler, D. C., and David, C. S., *Adv. Immun.*, **20**, 125–195 (1975).
16 Bach, F. H., Widmer, M. B., Bach, M. L., and Klein, J., *J. exp. Med.*, **136**, 1430–1444 (1972).
17 Streilein, J. W., and Billingham, R. E., *J. exp. Med.*, **126**, 455–473 (1967).
18 Staines, K. J., Guy, K., and Davies, D. A. L., *Eur. J. Immun.*, (in the press).
19 Lafferty, K. J., Cooley, M. A., Woolnough, J., and Walker, K. Z., *Science*, **188**, 259–261 (1975).
20 Sollinger, H. W., and Bach, F. H., *Nature*, **259**, 487–488 (1976).
21 Peck, A. B., and Bach, F. H., *Scand. J. Immun.*, **4**, 53–62 (1975).
22 Röllinghoff, M., and Wagner, H., *Eur. J. Immun.*, **5**, 875–879 (1975).

Possible mechanism for the biological action of lithium

IN spite of considerable early controversy it is now generally recognised that lithium has a favourable therapeutic effect on affective disorders and for some years lithium salts have been used with some degree of success in medical practice[1,2]. Curiously, in spite of the extreme simplicity of lithium as a drug, there is little understanding of its mode of action. In this article we outline a possible explanation.

Lithium as a drug is given in the form of a simple salt and
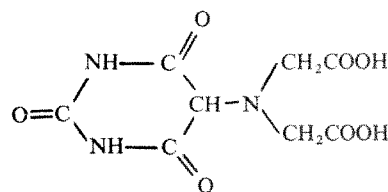
Fig. 1 Uramildiacetic acid

the active agent is undoubtedly the lithium cation. The dose level, around 1 mM, in the body, is extremely high and the most likely action of lithium is that it challenges one of the common biological cations Na^+, K^+, Mg^{2+} and Ca^{2+}. This explanation meets, however, with an immediate difficulty in that it is very hard to find an example of an organic ligand which can bind lithium with a stability constant as high as 10^3 and yet does not bind Mg^{2+} and Ca^{2+} with much higher affinity (Table 1). In other words, Table 1 seems to show that there is no ligand which lithium could bind in the presence of these cations. In this note we wish to show that the use of Table 1 in this fashion is completely misleading as to the possible action of lithium within the context of biological media, which amounts to saying that chemical affinities measured by absolute values of stoichiometric stability constants do not give the true tendency for preferential binding of a certain metal ion if other potential complexing agents are also present and that the free metal ion content of the compartment under discussion must be known. A better approach is to use 'conditional' stability constants,which are constants valid for the medium in which the reaction is taking place. The discussion then has much in common with that used in analytical chemistry where less stable metal complexes are preferentially formed by a metal M_1 with a ligand L_2 by 'removal' of other metals M_2 by 'sequestering' or 'masking' agents L_1 which bind M_2 differentially more strongly, that is, under the conditions of the experiment $K_{M_2L_1} \gg K_{M_1L_1}$ such that $K'_{M_1L_2} > K'_{M_2L_2}$.

We start the analysis of the binding of the five cations Li^+, Na^+, K^+, Mg^{2+}, Ca^{2+} to different ligands by using as a convenient measure of the uptake tendency of a certain ion by a certain ligand the product K'_{ML} [M], where K'_{ML} is the

Fig. 2 Conditional constants of the magnesium and lithium complexes of uramildiacetic acid ($\log K'_{ML}$, at $pH = 7$) as a function of the concentration of free ATP^{4-}.
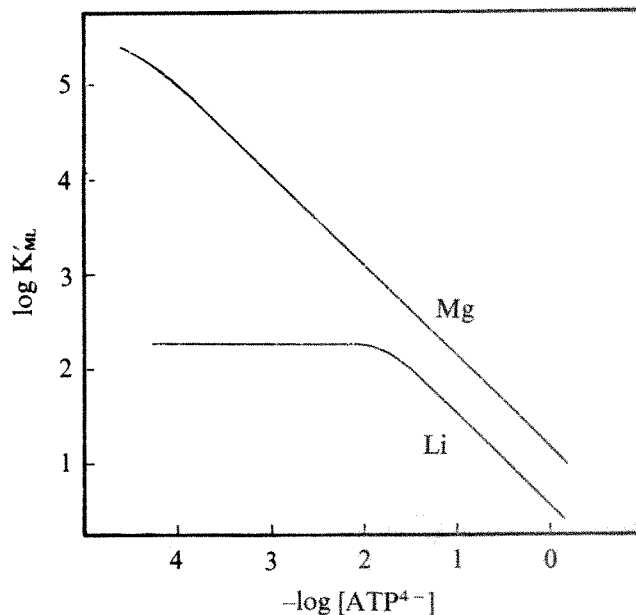
Table 1 Stability constants* (log K) of complexes of alkali and alkaline earth complexes in aqueous solution[7]

Ligand	Li$^+$	Na$^+$	K$^+$	Mg^{2+}	Ca^{2+}	$T(°C)$ µ†
OH$^-$	0.3	−0.5	—	2.5	1.3	25/0
HPO$_4^{2+}$	0.5	0.1	—	2.5	(1.5)	25/0
P$_2$O$_7^-$	2.39	1.0	0.8	5.41	4.95	25/1
P$_3$O$_{10}^{5-}$	2.8	1.7	1.3	5.83	5.44	25/1
ADP^{3-}	—	0.65	0.67	3.23	2.81	25/0.2
ATP^{4-}	1.7	1.2	0.9	4.43	3.92	25/0.3 or 0.1
Uramildiacetate	4.9	2.72	1.23	8.12	8.31	20/0.1
o-Hydroxyphenyliminodiacetate	2.2	1.0	—	6.86	6.27	20/0.1
o-Carboxyphenyliminodiacetate	2.05	0.89	—	3.91	5.06	20/0.1
Cyclohexyliminodiacetate	1.74	0.90	—	3.46	3.34	20/0.1
2-Hydroxycyclohexyliminodiacetate	2.19	0.75	—	4.27	5.19	20/0.1
o-Sulphophenyliminodiacetate	2.26	0.98	—	2.68	4.57	20/0.1
Iminodiacetate	0.96	0.36	—	2.94	2.59	20/0.1
Methyliminodiacetate	1.20	0.61	—	3.44	3.75	20/0.1
Nitrilotiracetate	2.51	1.32	0 6	5.41	6.41	20/0.1
EDTA^{4-}	2.79	1.66	0·.96	8.69	10.70	20/0.1
MethylEDTA^{4-}	3.43	2.24	—	10.29	11.47	30/0.1
EGTA^{4-}	1.2	1.4	—	5.41	11.0	25/1.5 or 0.1
DCTA^{4-} (ref. 4)	4.13	2.70	—	10.35	11.20	30/0.1
Cryptate 112 (ref. 5)	4.30	2.80	<2	—	2.80	25/0.1

†µ, Ionic strength.
*Source: *Stability Constants* (Chem. Soc. Sp. Publs 17 and 25, London 1964, 1971).

'conditional' stability constant of the complex formed between that ion and ligand concerned and [M] is the concentration of metal ion in solution in the compartment[3].

In the case of prolonged lithium therapy, the concentrations of the common alkali- and alkaline-earth metals in the cytoplasm of cells will be close to those given in Table 2. In the cells there are not many good ligands of comparable concentration but, strikingly, adenosine triphosphate (ATP) is often present in a very comparable concentration range. In some vesicles, such as the adrenal chromaffin granule, the ATP concentrations may reach even higher values; typical results are Ca^{2+} — 14 mM, Mg^{2+} — 6 mM, K$^+$ = 20 mM ATP — 100 mM (ref. 4 and A. Daniels and R. J. P. W., unpublished). Leaving such special regions of the cell aside for the moment, we see that inside the cytoplasm of cells generally lithium will be selected preferentially to sodium and potassium if log K$'_{Li}$ is greater than log K$'_{Na}$ and log K$'_{K}$ by a factor of 2. Several ligands meet this requirement. Lithium will be selected preferentially to calcium by any ligand which has a conditional stability constant which is not more than 3 log units smaller for the lithium complex and even up to a difference of 5 log units there will be appreciable competition of lithium for calcium sites. This is easily achieved by many ligands which have log K$'_{Li}$ of about 3 (that is conditional constant at pH = 7). Thus the calcium ion in cells will not prevent lithium binding. Magnesium poses a greater problem for it is present in equal or greater concentration than lithium and we must now examine competition between ligands for magnesium and lithium.

Let us take the situation in the cytoplasm when Mg^{2+} (total) < ATP (total) and examine the effect of this background of relatively high ATP concentration upon the conditional stability constant of lithium and magnesium complexes with some chosen ligands (Table 1).

One of the most stable lithium complexes known is that of

uramildiacetic acid ligand (Fig. 1)—which has been studied in detail by one of us[6,7].

The values given in Table 2 were used to calculate 'conditional' stability constants for the magnesium and the lithium complexes of uramildiacetic in the presence of 10$^{-2.5}$ M free ATP. The results are presented graphically in Fig. 2 and it can be seen that for the conditions prevalent inside cells, as described above, although the magnesium complex remains the most stable of the two, lithium is likely to interfere and compete with magnesium to some 20–30% for its binding sites.

It is perhaps worthwhile to point out that the choice of uramildiacetic acid is not arbitrary, as there are obvious analogies between uramildiacetic acid and some compounds from which it actually derives, for example, barbituric acid and uric acid, curiously enough associated with lithium therapy from the very beginning. Hence Lipowitz and Garrod[8] refer to the combination of lithium with uric acid promoting the dissolution of urate deposits in cartilages and, more recently, Margulies[9] used lithium in combination with barbiturates obtaining favour-

Table 2 Concentration of common alkali earth metals and of ATP in cytoplasm of cells

Species	Concentration (M)
Li$^+$	10^{-3}
Na$^+$	10^{-2}
K$^+$	10^{-1}
Mg^{2+}	10^{-3} — 10$^{-2.5}$
Ca^{2+}	10^{-7}
ATP	10$^{-3.0}$ — 10$^{-2.5}$

Fig. 3 Conditional constants of the magnesium and lithium complexes of o-carboxiphenyliminodiacetic acid (log K$'_{ML}$) at pH = 7 as a function of the concentration of free ATP^{4-}.
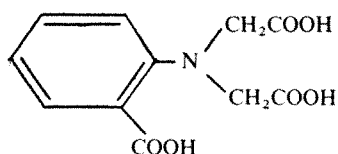
Fig. 4 Ortho-carboxyphenyliminodiacetic acid.

able results in the treatment of manic states. Protective effects on urea toxicity in urine from manic–depressive patients injected to guinea pigs was also reported by Case[10].

Nevertheless, uramildiacetic acid as such does not occur in biological systems and the conclusion to be derived from the result expressed in Fig. 1 is that chelating sites of liganding atoms (O⁻, N, O⁻, O⁻) are likely to exhibit increased affinity for lithium and, in the presence of alkaline-earth metal sequestering agents, this metal may compete to a greater or lesser extent with Mg^{2+} and Ca^{2+} for the above mentioned sites.

As a matter of fact it is not necessary that the ligands have a particularly high affinity for lithium, if it is high enough to give appreciable complex formation. What is more important is that the gap of stability between the lithium and the magnesium complexes is made as small as possible while keeping the values of the constants above 2 log units. In Fig. 3 we present the values obtained for the lithium and magnesium complexes of o-carboxyphenyliminodiacetate (Fig. 4), which has a coordinating site (O⁻, N, O⁻, O⁻) similar to that of uramildiacetic acid, although it is not such a powerful complexing agent[11]. It can be seen that for concentrations of free ATP of the order of $10^{-2.5}$ M the lithium and magnesium complexes are of equal 'effective' stability and for higher concentrations of ATP the lithium complex actually becomes stronger than the magnesium complex; lithium will then be taken up preferentially by this (O⁻, N, O⁻, O⁻) ligand, although log K'_{Li} at $pH = 7$ is not sufficiently high to ensure the predominance of the complex species.

Other similar examples can be given and although the ligands referred to above do not exist *in vivo*, it is likely that analogous or even more adequate coordinating sites may have developed. The variety of sites provided *in vivo* for the uptake of the various elements and the characteristics of those for calcium and magnesium make this idea more than a simple guess or prediction, see, for example, concanavalin[12].

If we consider proteins as ligands, it is obvious that sites of the kind (O⁻, N, O⁻, O⁻) could have evolved, that they would bind Mg^{2+}, but that this binding would be considerably affected by the concentration of ATP and, finally, that in this complicated situation there could be a further weakening of binding through the influence of added lithium. Now, while a situation in which $(Mg^{2+}) = 10^{-3}$ M and $(ATP) = 10^{-2.5}$ M might be thought to be somewhat artificial in the cytoplasm of cells in general, just such a situation arises in the internal fluids of the adrenal chromaffin granule and it could occur in many other vesicles, for example, in the transmitter vesicles of nerve cells. Thus it could well be that Li^+ is indeed challenging Mg^{2+} and so affecting metabolic rates as diverse as the action of glycolysis to the rates of many ATP-driven transfer proteins.

J. J. R. Frausto da Silva*

Centro de Química Estrutural,
Instituto Superior Tecnic, Lisbon, Portugal

R. J. P. Williams

Wadham College and Inorganic Chemistry Laboratory,
University of Oxford,
Oxford OX1 3QR, UK

1 *Lithium* (edit. by Gershon, S., and Shopsin, B.), (Plenum, New York, 1973).
2 *Lithium Research and Therapy* (edit. by Johnson, F. N.), (Academic, London, 1975).
3 Frausto da Silva, J. J. R., and Williams, R. J. P., *Structure and Bonding* (in the press).
4 Borowitz, J. L., Fuwa, K., and Wiener, N., *Nature*, 205, 42–43 (1965).
5 Lehn, J. M., and Sauvage, J., *J. chem. Soc. chem. Comm.*, 440–441 (1971).
6 Irving, H. M. N. H., and Frausto da Silva, J. J. R., *J. Chem. Soc.*, 448, 458 (1963).
7 Frausto da Silva, J. J. R., Candida, M., and Vaz, T. A., *Rev. Port Quimica*, 14, 102–109 (1972).
8 Kline, N. S., in *Lithium* (edit. by Gershon, S., and Shopsin, B.), (Plenum, New York, 1973).
9 Margulies, M., *Med. J. Australia*, 41, 137–141 (1955).
10 Case, J. T. T., *Med. J. Australia*, 36, 349–352 (1945).
11 Birch, N. J., *Lancet*, i, 1384 (1945).
12 Hardman, K. B., in *Metal Ions and Biological Systems* (edit. by Dhar, S. K.), 103–123 (Plenum, New York, 1973).

β-Endorphin as a potent analgesic by intravenous injection

An untriakontapeptide[1] with significant opiate activity from camel pituitary glands was shown to have an amino acid sequence identical to the carboxyl-terminal 31 amino acids of β-lipotropins[2,3]. The peptide, designated as β-endorphin, has recently been synthesised[4], and when administered directly into the brain, it was found[5] to be considerably more potent than morphine in molar basis, and its actions were blocked by naloxone. Other peptides related to β-lipotropins have also been shown to possess *in vitro* opiate-like activity[6-9]. We report here the analgesic effects of β-endorphin when it is injected intravenously.

Male ICR mice weighing 25–30 g (Simonson Laboratories) were used in all the experiments. β-Endorphin was synthesised as described previously[4]. Naloxone HCl was a gift from Endo Laboratories and morphine sulphate was purchased from Mallinckrodt Chemical Works. The antinociceptive or analgesic

Fig. 1 Antinociceptive effects following the intravenous injection of β-endorphin in (a) the tail-flick test and (b) hotplate test, and its blockage by naloxone. Mice, 6–7 per group, were injected intravenously with 8.2 (△), 14.5 (□), and 20.1 (○) mg kg⁻¹ of β-endorphin at 0 time. Naloxone HCl (1 mg kg⁻¹) (●) was injected subcutaneously 5 min before the intravenous injection of β-endorphin. 'Percentage' antinociceptic effect was calculated[12] as $[(T_1 - T_0)/(T_2 - T_0)] \times 100$, where a control latency (T_0) was obtained from the mean of two latencies determined before drug injection: test latencies (T_1) were determined at various times after injection for each animal; the cutoff times (T_2) for the tail-flick test and hotplate test were 7 and 60 s, respectively. The vertical bars indicate the s.e.m.

Time after β-endorphin intravenous injection (min)
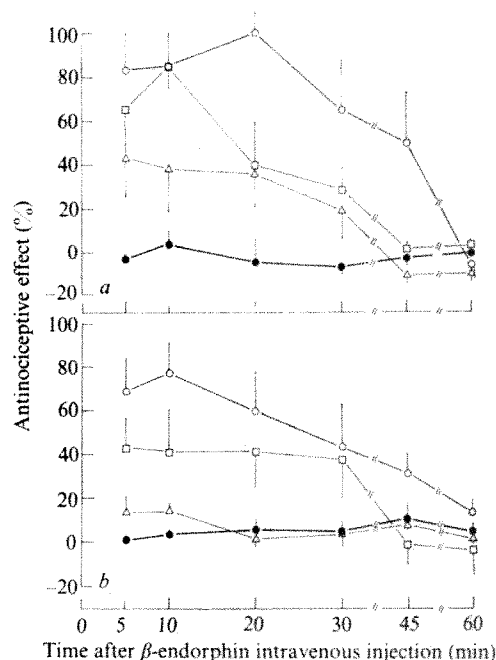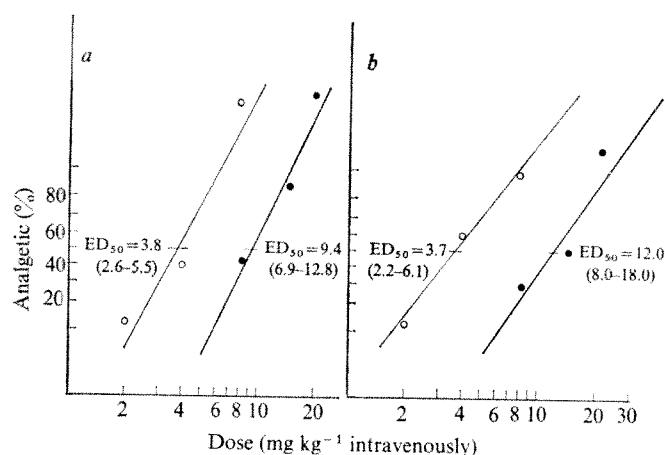
Fig. 2 Comparison of antinociceptive effects of β-endorphin (●) and morphine (○) in tail-flick test (*a*) and hotplate test (*b*). After control latency was obtained before the drug injection, groups of mice, 6–10 in each, were injected intravenously with various doses of β-endorphin or morphine sulphate and the antinociceptive effect was determined 10 min after drug injection. The values in the figure are median antinociceptive dose (ED_{50}) in mg kg^{-1} and the 95% confidence limits are shown in the parentheses. When the ED_{50} values are expressed on the molar basis, the ED_{50} of β-endorphin is 2.7 μmol kg^{-1} and of morphine 11.4 μmol kg^{-1} in the tail-flick test. The ED_{50} of β-endorphin is 3.5 μmol kg^{-1} and of morphine 11.1 μmol kg^{-1} in the hotplate test.

properties of β-endorphin and morphine were assessed by both the hotplate[10] and the tail-flick[11] methods. β-Endorphin and morphine sulphate were injected intravenously into the tail vein in a volume of 0.01 ml per g body weight. The extent of analgesia or antinociceptive response was expressed as 'percentage antinociceptive effect' (see ref. 12).

With a twofold increase in latency of reaction times as a quantal index of inhibition, the median antinociceptive dose (ED_{50}) and 95% confidence limits were calculated according to the method of Litchfield and Wilcoxon[13].

β-Endorphin administered at intravenous doses of 8.2, 14.5 and 20.1 mg kg^{-1} produced a dose-related inhibition of both the tail-flick and the hotplate response of mice to nociceptive stimuli (Fig. 1). These effects lasted 30–60 min, depending on the dose used, and seemed to be shorter than the duration of β-endorphin when applied centrally[5]. Pretreatment of mice with naloxone HCl (1 mg kg^{-1}, subcutaneously) 5 min before the injection of 20.1 mg kg^{-1} of β-endorphin completely eliminated the inhibitory effect of β-endorphin on the tail-flick and hotplate responses. Thus, in two tests for antinociceptive activity, β-endorphin injected intravenously was as effective as when it was administered directly into the brain.

The percentage of mice which were analgetic after intravenous injection of various doses of β-endorphin and morphine, and their ED_{50} values with 95% confidence limits are given in Fig. 2. The ED_{50} values expressed in mg kg^{-1} of β-endorphin were about 2–3 times higher than that of morphine. When the potency was compared on a molar basis, β-endorphin was ~ 3–4 times more potent than morphine. We have previously shown that β-endorphin was ~ 18–33 times more potent than morphine when applied centrally[5]. Since the molecular weight of β-endorphin is ~ 10 times higher than morphine, it is likely that the lower potency ratio of β-endorphin to morphine when they are applied peripherally is due to the relatively poorer penetration of β-endorphin into the brain. In addition, a different rate of degradation after peripheral administration is a possible factor.

We have shown that β-endorphin can produce analgesia when administered intravenously and the analgesic effect can be eliminated by pretreatment with naloxone. This work was

Hormone Research Laboratory and
Langley Porter Neuropsychiatric Institute,
Department of Pharmacology,
University of California, San Francisco, California 94143

1 Li, C. H., and Chung, D., *Proc. natn. Acad. Sci. U.S.A.*, 73, 1145–1148 (1976).
2 Li, C. H., Barnafi, L., Chrétien, M., and Chung, D., *Nature*, 208, 1093–1094 (1965).
3 Li, C. H., and Chung, D., *Nature*, 260, 622–624 (1976).
4 Li, C. H., Lemaire, S., Yamashiro, D., and Doneen, B. A., *Biochem. biophys. Res. Commun.*, 71, 19–25 (1976).
5 Loh, H. H., Tseng, L. F., Wie, E., and Li, C. H., *Proc. natn. Acad. Sci. U.S.A.*, 73, 3895–3898 (1976).
6 Hughes, J., Smith, T. W., Kosterlitz, H. W., Fothergill, L. A., and Morris, H. R., *Nature*, 258, 577–579 (1975).
7 Bradbury, A. F., Smyth, D. G., Snell, C. R., Birdsell, N. J. M., and Hulme, E. C., *Nature*, 260, 793–795 (1976).
8 Lazarus, L. H., Ling, N., and Guillemin, R., *Proc. natn. Acad. Sci. U.S.A.*, 73, 2156–2159 (1976).
9 Gráf, L., Rónai, A., Bajusz, S., Cseh, G., and Székely, J. I., *FEBS Lett.*, 64, 181–184 (1976).
10 Antier, S., *Eur. J. Pharmac.*, 27, 1–4 (1974).
11 D'Amour, F. E., and Smith, D. L., *J. Pharmac. exp. Ther.*, 72, 74–79 (1941).
12 Harris, L. S., and Pierson, A. K., *J. Pharmac. exp. Ther.*, 143, 141–148 (1964).
13 Litchfield, J. T., and Wilcoxon, F., *J. Pharmac. exp. Ther.*, 96, 99–113 (1949).

Comparative study on analgesic effect of Met^5-enkephalin and related lipotropin fragments

Hughes *et al.*[1] have reported the isolation and structure of two pentapeptides from porcine brain with opiate agonist activity in isolated systems. The structure of one of these peptides, Met^5-enkephalin, is identical with the sequence of pituitary β-lipotropin (β-LPH) between residues 61–65 (refs 2–4). To prove the biological correlation of brain enkephalin and pituitary β-LPH, a series of lipotropin fragments, LPH-(61–69)-[5], LPH-(61–76)-[6] and LPH-(61–91)-peptides[7–10], have been isolated and shown to have opiate agonist activity *in vitro*. Only few and controversial data have been available so far on the analgesic effect *in vivo* of the above or similar lipotropin fragments. Enkephalins have recently been reported to induce analgesia *in vivo*[11,12]. Our preliminary data[5,13], however, seemed to contradict these observations, rather suggesting that some larger fragment(s) of β-LPH may have analgesic properties. We have therefore compared the analgesic effects of Met^5-enkephalin and some lipotropin fragments containing the complete structure of Met^5-enkephalin at their NH$_2$-terminus. The results show that the *in vivo* effect is a function of the length of the peptide chain, Met^5-enkephalin being the least and LPH-(61–91)-peptide the most potent. During the preparation of this paper we have become aware of the recent observation of Bradbury and coworkers[10,14] on a strong analgesic activity of LPH-(61–91)-peptide.

Met^5-enkephalin was synthesised as described previously[5]. LPH-(61–79)-peptide was obtained by plasmin cleavage of porcine β-LPH[15–17]. LPH-(61–91)-peptide was isolated from a crude adrenocorticotropic hormone preparation by high voltage paper electrophoresis at pH 6.5 (ref. 9). The *in vitro* opiate agonist activity of the fragments was determined on longitudinal muscle strip of guinea pig ileum[18]. For the *in vivo* assay the classical tail-flick technique was used according to D'Amour and Smith[19]. The tail of male CFY rats was exposed to a radiant heat and the latency of tail withdrawal measured automatically with the accuracy of 0.1 s. The cutoff time was 15 s. All the substances studied were dissolved in saline and administered intra-

cerebroventricularly as described by Chermat and Simon[20], in a volume of 20 μl.

In the *in vitro* assay the potency order was found to be as follows: LPH-(61–91)-peptide=LPH-(61–65)-peptide> LPH-(61–69)-peptide=LPH-(61–79)-peptide (Table 1). The 70–90% inhibitory effect of these fragments could be reversed by 100 nM naloxone.

Using the schedule described by Belluzzi *et al.*[11] for measurement the analgesic effect of peptides—that is, measuring the latency time every 2 min for 10 min—the prolongation of the reaction time was moderate for all the materials studied, including morphine. In addition, the effect of enkephalin was short lasting, subsiding 6–8 min after application (Fig. 1*a*).

This type of *in vivo* effect significantly differs from that caused by either morphine or larger lipotropin fragments since the maximal effect of the latter compounds did not develop within such a short period of time. As Fig. 1*b* shows, using another time schedule—that is, controlling the tail withdrawal reaction for 2 h—a marked analgesic activity was obtained for morphine, LPH-(61–91)- and LPH-(61–79)-peptides, but only a moderate and transitory prolongation of the tail-flick reaction could be achieved by Met⁵-enkephalin. The three stronger analgesic compounds were characterised by a slow onset of action which reached a plateau at 15–30 min after administration and lasted for 60–120 min. Their activity could be antagonised by subcutaneously administered naloxone at a dose of 2 mg kg⁻¹, 30 min before intracerebroventricular treatment.

The *in vitro* and *in vivo* opiate agonist activities of Met⁵-

Fig. 1 Time–response curves for the analgesic activities of morphine and β-LPH fragments administered intracerebroventricularly to rats. Abscissa: time elapsed before (−) and after intracerebroventricular administration of materials (at time 0). Latency time was measured either every 2 min (*a*) or using longer intervals between the measurements as indicated on the scale (*b*). □, Morphine 33 nmol per animal ($n = 20$); ○, LPH-(61–91)-peptide 0.8 nmol per animal ($n = 6$); ●, LPH-(61–65)-peptide 670 nmol per animal ($n = 23$); ×, saline 20 μl per animal ($n = 21$) for *a*; and □, morphine 33 nmol per animal ($n = 19$); ○, LPH-(61–91)-peptide 0.8 nmol per animal ($n = 6$); ▲, LPH-(61–79)-peptide 40 nmol per animal ($n = 5$); ●, LPH-(61–65)-peptide 670 nmol per animal ($n = 10$); ×, saline 20 μl per animal ($n = 16$) for *b*, respectively.
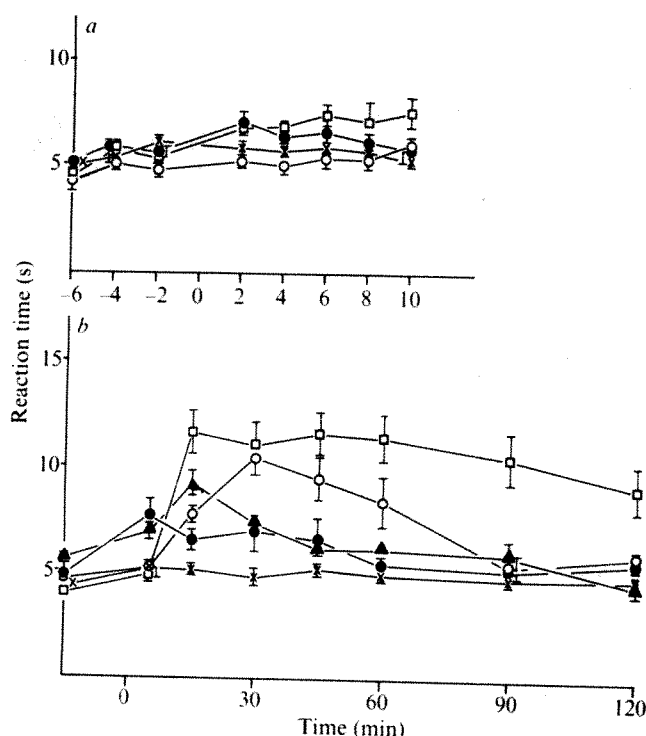
Table 1 Opiate agonist potencies of porcine β-LPH fragments in longitudinal muscle strip of guinea pig ileum

Fragment	Relative agonist potencies*	Statistical evaluation†
LPH-(61–65)-peptide	1 ($n = 4$)	—
LPH-(61–69)-peptide	0.51 ± 0.12 ($n = 4$)‡	$P < 0.05$
LPH-(61–79)-peptide§	0.42 ± 0.06 ($n = 7$)	$P < 0.005$
LPH-(61–91)-peptide§	1.37 ± 0.16 ($n = 4$)	NS

*Potency ratios were calculated from ID_{50} values. The dose–response curves for the peptides were parallel.

†The statistical comparison was based on the ID_{50} values, using the unpaired *t* test.

‡Mean ± s.e.m. values are listed; the numbers of experiments are in parentheses.

§Guillemin *et al.*[6] and Li and Chung[7] have proposed to designate LPH-(61–76)-, LPH-(61–77)- and LPH-(61–91)-peptides as α, γ and β endorphins, respectively. Considering that LPH-(61–79)-peptide also occurs in the pituitary gland (unpublished results), and also that it is closely similar to γ endorphin, we propose to designate it as δ endorphin.

enkephalin (Table 1, Fig. 1) indicate that the molecule contains all the essential residues for the interaction with the opiate receptors. The relative inefficiency and short duration of the *in vivo* effect induced by Met⁵-enkephalin may be due to its rapid degradation by brain enzymes, as also suggested by Belluzzi *et al.*[11]. The more pronounced and prolonged analgesic effect obtained with LPH-(61–79)-peptide strongly supports this view (Fig. 1*b*). LPH-(61–79)-peptide is less potent than Met⁵-enkephalin on longitudinal muscle strip of guinea pig ileum. Consequently the function of the COOH-terminal portion (residues 66–79) of the molecule in increasing the *in vivo* analgesic effect of the peptide may be the protection of the active site (residues 61–65) against the proteolytic enzymes of the brain. A similar protecting role has been attributed to the COOH-terminal portion of adrenocorticotropic hormone (residues 25–39)[21,22]. Among the fragments studied, LPH-(61–91)-peptide showed the highest opiate agonist activity *in vitro* (Table 1). This finding, together with the extremely high *in vivo* analgesic effect of this fragment (at least 50 times higher than that of morphine), suggests that additional receptor binding site(s) may be present within residues 80–91 of the molecule.

Our *in vivo* results and also their explanation seem to be in agreement with the recent data of Bradbury *et al.*[10] on the *in vitro* receptor-binding properties of similar fragments.

We thank Dr T. Lang for encouragement, and Drs J. Borsy, G. Cseh and L. Tardos for discussions.

L. Graf
J. I. Szekely
A. Z. Ronai
Zsuzsa Dunai-Kovacs
S. Bajusz

Research Institute for Pharmaceutical Chemistry,
H-1325 Budapest,
PO Box 82, Hungary

1 Hughes, J., *et al.*, *Nature*, **259**, 577–579 (1975).
2 Li, C. H., Barnafi, L., Chrétien, M., and Chung, D., *Nature*, **208**, 1093–1094 (1965).
3 Gráf, L., Barát, E., Cseh, G., and Sajgó, M., *Biochim. biophys. Acta*, **229**, 276–278 (1971).
4 Cseh, G., Barát, E., Patthy, A., and Gráf, L., *FEBS Lett.*, **21**, 344–346 (1972).
5 Gráf, L., Rónai, A. Z., Bajusz, S., Cseh, G., and Székely, J. I., *FEBS Lett.*, **64**, 181–184 (1976).
6 Guillemin, R., Ling, N., and Burgus, R., *C. r. hebd. Séanc. Acad. Sci. Paris*, **D282**, 783–785 (1976).
7 Li, C. H., and Chung, D., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 1145–1148 (1976).
8 Bradbury, A. F., Smyth, D. G., and Snell, C. R., *Nature*, **260**, 165–167 (1976).
9 Gráf, L., Barát, E. and Patthy, A., *Acta biochim. biophys. Acad. Sci. Hung.*, **11**, 121–122 (1976).
10 Bradbury, A. F., Smyth, D. G., Snell, C. R., Birdsall, N. J. M., and Hulme, E. C., *Nature*, **260**, 793–795 (1976).
11 Belluzzi, J. D., *et al.*, *Nature*, **260**, 625–626 (1976).
12 Büscher, H. H., *et al.*, *Nature*, **261**, 423–425 (1976).
13 Rónai, A. Z., Székely, J. I., Gráf, L., Dunai-Kovács, Zs., and Bajusz, S., *Life Sci.* (in the press).

14 Feldberg, W. A., and Smith, D. G., J. Physiol., Lond. (in the press).
15 Gráf, L., and Cseh, G., Acta biochim. biophys. Acad. Sci. Hung., 3, 175–177 (1968).
16 Gráf, L., and Li, C. H., Biochem. biophys. Res. Commun., 53, 1304–1309 (1973).
17 Gráf, L., Acta biochim. biophys. Acad. Sci. Hung. (in the press).
18 Paton, W. D. M., and Vizi, E. S., Br. J. Pharmac., 35, 10–28 (1969).
19 D' Amour, F. E., and Smith, D. L., J. Pharm., 72, 74–79 (1941).
20 Chermat, R., and Simon, P., J. Pharmac., Paris, 6, 489–492 (1975).
21 Schwyzer, R., Schiller, P., Seelig, S., and Sayers, G., FEBS Lett., 19, 229–231 (1971).
22 Bennett, H. P. J., Bullock, G., Lowry, P. J., McMartin, C., and Peters, J., Biochem. J., 138, 185–194 (1974).

Selective activation of the mesocortical DA system by stress

IT is now well established that an important system of dopaminergic (DA) neurones innervates various parts of the cerebral cortex in the rat and other species[1–3]. In contrast to noradrenergic (NA) terminals which are widely distributed in this structure, the DA terminals are mainly confined to deep layers, particularly in the frontal, the cingular and the entorhinal areas[4,5]. The results of lesion studies demonstrated that the terminal endings in the frontal cortex originate from the A10 group of DA cell bodies localised in the mesencephalon[6–8]. This group also contains the cell bodies of the classical mesolimbic DA system projecting to the tuberculum olfactorium, the nucleus accumbens, the nucleus of the stria terminalis, and the amygdala[9]. The DA terminals found in the cingulate and entorhinal areas of cortex may originate mainly from the A9 group of DA neurones[6–8]. This group gives rise to the well known nigrostriatal DA system which is implicated in extrapyramidal processes. Its degeneration is in part responsible for some of the symptoms seen in Parkinsonian patients. Little is yet known about the functions of the mesocortical and mesolimbic DA pathways. Electrocoagulation or 6-hydroxydopamine (6-OHDA)-induced lesions of the ventral tegmental area, containing the A10 group, produce a syndrome characterised by "locomotor hyperactivity, serious impairment in tests requiring inhibition of a previously learned response, facilitation of approach learning and of active avoidance and hypoemotivity"[10,11]. Various workers have suggested that the antipsychotic effects of neuroleptics are in part related to the blockade of postsynaptic DA receptors localised in areas innervated by the mesolimbic and mesocortical DA systems[12,13]. It seems important to establish whether neurones of these two DA systems correspond to an homogeneous population of cells with similar functional characteristics. We have therefore explored this problem in the rat by examining the reactivity of the mesocortical and mesolimbic DA pathways as well as that of nigrostriatal DA system to stress induced by electric foot shocks. Our results suggest that the mesocortical DA system is selectively activated by this stress.

Groups of male Charles River rats of the Sprague–Dawley strain, kept in a controlled environment (24 °C, 60% humidity, alternate cycles of 12 h light—0700–1900—and 12 h darkness 1900–0700) and receiving food and water ad libitum were injected at 1400 with α-methylparatyrosine (200 mg kg⁻¹ intraperitoneally), an inhibitor of catecholamine (CA) synthesis. Some of them then received electric foot shocks 10 min later, using a procedure previously described[14]. Others were used as controls. All animals were killed 30 min after administration of the synthesis inhibitor. Changes in CA concentrations were estimated in the medial part of the frontal cortex, the nucleus accumbens and the tuberculum olfactorium, and in the striatum. Since the two first areas are relatively poor in DA or consist of only small amounts of tissue, they were dissected precisely from frozen frontal slices (500 μm thick) using the microdisk technique originally described by Palkovits[15]. Particular care was taken to punch out areas rich in DA fibres

and 3H-DA uptake sites in the medial part of the frontal cortex[8] between planes 9,800 and 12,800 μm; the beginning of the striatum at 9,800 μm according to the atlas of König and Klippel[16] was used as a landmark. The nucleus accumbens was also punched out with a tube sharpened at the extremity (0.8 mm diameter) from two slices delimited by the 9,800- and 8,800-μm planes. Pooled microdisks of tissues corresponding to each area were sonicated in 80 μl of 0.1 N perchloric acid and 0.01 N thioglycollic acid solution. After centrifugation, DA and NA were isolated from the supernatant on alumina microcolumns[17], and measured using the sensitive radioenzymatic microassay of Gauchy et al.[17]. Pellets were dissolved in 0.1 N NaOH for protein estimations[18]. Since the tuberculum olfactorium and the striatum contain large amounts of DA, they were dissected at 4 °C and the amine was isolated by ion-exchange chromatography on Amberlite CG50 and estimated spectrofluorimetrically[14]. To compare the results obtained by the two procedures, in one experiment, DA was estimated radioenzymatically in two microdisks punched out in the mediolateral part of the striatum from two slices (500 μm thick) delimited by the 8,500- and 7,500-μm planes.

Changes in the rate of disappearance of CA after synthesis inhibition provide reliable indications of modifications in the rate of amine utilisation. The electric foot shock stress did not affect the rate of DA utilisation either in the whole striatum or in its mediolateral part, since DA levels were identical in control and stressed animals (Fig. 1.) Similar observations were made in the tuberculum olfactorium. In the nucleus accumbens, however, which is another structure innervated by the mesolimbic DA system, the utilisation of DA was accelerated as indicated by the 25% decrease in DA levels in tissues from stressed animals when compared with those of respective controls (Fig. 1). The stress effect was much more striking in the
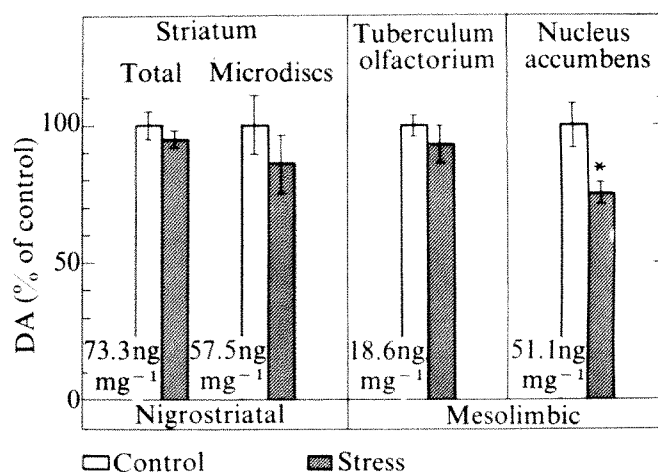
Fig. 1 Effect of 20-min electric foot shock stress on the utilisation of DA in structures innervated by the nigrostriatal or the mesolimbic DA pathways. DA was estimated spectrofluorimetrically (total striatum, tuberculum olfactorium) or radioenzymatically (striatal microdisks, nucleus accumbens) 30 min after the injection of α-methylparatyrosine (200 mg kg⁻¹, intraperitoneally) in tissues of control and stressed rats. The stress was applied 10 min after the injection of inhibitor. In the radioenzymatic assay, α-methyldopamine is measured concomitantly with DA. However, α-methyldopamine formed from α-methylparatyrosine represented less than 10% of DA levels 30 min after the inhibitor injection (A. Chéramy, personal communication). Numbers in bars represent absolute mean vaues of DA (ng per mg protein). Data expressed as % of respective control values are the mean ± s.e.m. of results obtained with groups of 6 rats. P < 0.02 when compared with respective controls. DA levels in animals not treated with α-methylparatryrosine were for total striatum, striatal microdisks tuberculum olfactorium and nucleus accumbens: 90.0 ± 6.0 ng mg⁻¹, 75.0 ± 5.0 ng mg⁻¹, 22.3 ± 2.1 ng mg⁻¹ and 79.1 ± 7.0 ng⁻¹ respectively.
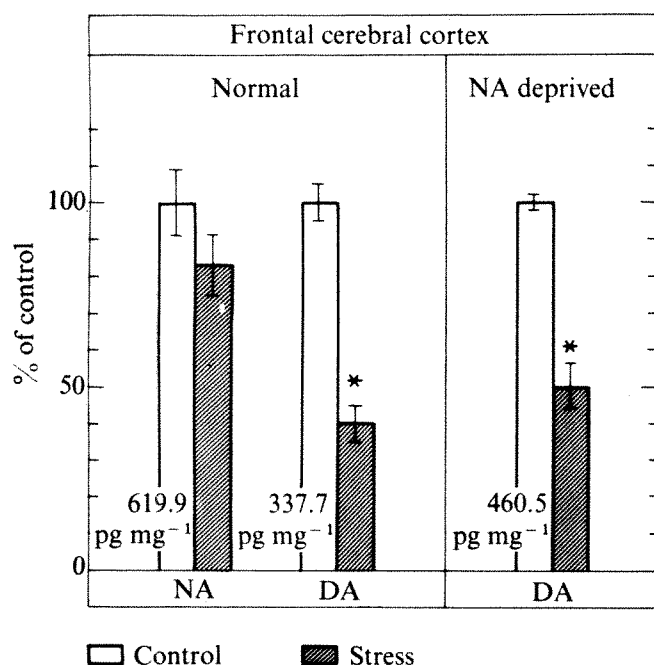
Fig. 2 Effects of a 20-min electric foot shock stress on the utilisation of DA in the frontal cerebral cortex of normal rats and of animals deprived of cortical NA innervation (NA deprived). Selective degeneration of ascending NA pathways was induced by local 6-OHDA injections 12 weeks before the stress experiment. All rats were submitted to the 20-min stress situation 10 min after the injection of α-methylparatyrosine (200 mg kg^{-1} intraperitoneally). DA and NA levels were estimated radioenzymatically in the frontal cerebral cortex 30 min after the inhibitor injection (numbers in bars represent the absolute mean values of NA and DA expressed as pg per mg protein). Data expressed as % of respective control values are the mean ± s.e.m. of results obtained with groups of six rats. In rats deprived of noradrenergic innervation, cortical NA levels were less than 10% those of controls. $P < 0.001$. Cortical DA and NA levels in normal rats not treated with α-methylparatyrosine were 720 ± 30 pg mg^{-1} and 950 ± 100 pg mg^{-1}.

frontal cerebral cortex, in which a 60% reduction in DA levels was observed in animals submitted to electric foot shocks (Fig. 2). The short period of stress did not affect the rate of NA utilisation in this structure.

The effect of stress on the mesocortical DA neurones was further studied in rats deprived of a NA innervation to the cerebral cortex. For this complementary experiment, the ascending NA systems were destroyed by a local microinjection of 6-OHDA (2 μg in 1 μl of isotonic saline solution) made laterally to the pedunculus cerebellaris superior[1], 12 weeks before the stress experiment. Two groups of rats were injected as before with α-methylparatyrosine and one of them was exposed to the stress situation. As indicated in Fig. 2, the rate of DA utilisation was accelerated in the stressed rats in a similar manner to that observed in unlesioned animals. As expected, NA could no longer be detected in the frontal cerebral cortex of the lesioned rats. It thus seemed that the stress-induced activation of the mesocortical DA neurones was unrelated to the ascending NA systems.

In previous studies[14], we reported that an electric foot shock stress of long duration accelerated both the utilisation and the turnover of NA in various structures of the brain; the effect being particularly striking in the brainstem. We assumed that these changes in NA metabolism were related to modifications in neuronal activity. The stress of long duration (180 min) did not influence DA utilisation in striatal DA terminals[14]. It is thus not surprising that the 20-min stress used in the present study did not modify the activity of the nigrostriatal DA neurones. In fact, other

stress situations such as exposure to noise or photic stimuli, intense muscular exercises, or exposure to cold, were found to have no effects on DA metabolism in the striatum. In the present study, it was interesting to note that two structures belonging to the mesolimbic DA system did not react similarly to the 20-min stress situation. Indeed DA utilisation was accelerated significantly in the nucleus accumbens but not in the tuberculum olfactorium. This demonstrates that the various DA neurones of the mesolimbic system do not behave similarly. We have already indicated that the DA terminals localised in the medial part of the frontal cortex also originate from the A10 group of DA cell bodies. The dramatic effect of electric foot shocks on the activity of these mesocortical neurones further revealed the heterogeneity of the A10 group. The very pronounced and selective acceleration of DA utilisation in these neurones indicates that the mesolimbic and mesocortical DA systems can be distinguished not only on anatomical grounds but also from a functional point of view. From these results, it is doubtful that the DA terminals present in the various limbic structures and in the frontal cortex originate from the same DA cell bodies. Recent results on the effects of chronic neuroleptic treatments on the synthesis of DA in structures innervated by the mesolimbic and mesocortical DA pathways support the above statements[19]. Indeed, the time curves of the effect of neuroleptics on DA synthesis in the frontal cerebral cortex and in the tuberculum olfactorium + nucleus accumbens were different.

In an earlier study, Bliss and Ailion[20] observed an increase in the content of homovanillic acid (a major DA metabolite) in the whole brains of rats or mice submitted to severely stressful experiments. They suggested that DA neurones could be involved in the cerebral circuit of emotionality. In this context, the high and selective sensitivity of the mesocortical DA neurones to electric foot shock stress revealed in our study is of particular interest. This DA system originates from the "limbic midbrain area" and projects to the frontal cerebral cortex, connected directly and indirectly to the limbic structures. It may be that this mesocortical DA pathway has a critical role in the regulation of emotional states.

A. M. Thierry
J. P. Tassin
G. Blanc
J. Glowinski

Group NB—INSERM U.114,
Collège de France,
11. place Marcelin Berthelot,
75231—Paris cedex 5, France

[1] Thierry, A. M., Blanc, G., Sobel, A., Stinus, L., and Glowinski, J., Science, 182, 499–501 (1973).
[2] Thierry, A. M., Hirsch, J. C., Tassin, J. P., Blanc, G., and Glowinski, J., Brain Res., 79, 77–88 (1974).
[3] Divac, I., Lindvall, O., Björklund, A., and Passingham, R. E., Expl Brain Res., suppl. vol. 23, 58 (1975).
[4] Hökfelt, T., Fuxe, K., Johansson, O., and Ljungdahl, Å., Eur. J. Pharmac., 25, 108–112 (1974).
[5] Berger, B., Thierry, A. M., Tassin, J. P., and Moyne, M.A., Brain Res., 106, 133–145 (1976).
[5] Fuxe, K., Hökfelt, T., Johanson, O., Jonsson, G., Lidbrink, P., and Ljungdahl, A., Brain Res., 82, 349–355 (1974).
[7] Lindvall, O., Björklund, A., Moore, R. Y., and Stenevi, U., Brain Res., 81, 332–337 (1974).
[8] Tassin, J. P., Blanc, G., Stinus, L., Berger, B., Glowinski, J., and Thierry, A.M., in Advances in Biology and Medicine, 69 (edit. by Levi, G., Battistin, L., and Lajtha, A.), (Plenum, New York, in the press).
[9] Ungerstedt, U., Acta physiol. scand., suppl. 367 (1971).
[10] Le Moal, M., Cardo, B., and Stinus, L., Physiol. Behav. 4, 567–573 (1969).
[11] Galey, D., Simon, H., and Le Moal, M., Brain Res. (in the press).
[12] Anden, N. E., Butcher, S. G., Corrodi, H., Fuxe, K., and Ungerstedt, U., Eur. J. Pharmac., 11, 303–314 (1970).
[13] Scatton, B., Thierry, A. M., Glowinski, J., and Julou, L., Brain Res., 88, 389–393 (1975).
[14] Thierry, A. M., Javoy, F., Glowinski, J., and Kety, S. S., J. Pharmac. exp. Ther., 163, 163–171 (1968).

15 Palkovits, M., *Brain Res.*, 59, 449–450 (1973).
16 König, J. F. R., and Klippel, R. A., *The Rat Brain: A stereotaxic Atlas of the Forebrain and Lower Parts of the Brain Stem* (Williams and Wilkins, Baltimore, 1963).
17 Gauchy, C., Tassin, J. P., Glowinski, J., and Chéramy, A., *J. Neurochem.*, 26, 471–480 (1976).
18 Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J., *J. biol. Chem.*, 193, 265–275 (1951).
19 Scatton, B., Glowinski, J., and Julou, L., *Brain Res.* (in the press).
20 Bliss, E. L., and Ailion, J., *Life Sci.*, 10, 1161–1169 (1971).

Lesion of striatal neurones with kainic acid provides a model for Huntington's chorea

THE symptoms of Huntington's chorea, an hereditary movement disorder, result from degeneration of neurones primarily in the basal ganglia[1]. Several neurochemical abnormalities have been identified in the brains of patients dying with this disorder[2–5], but no animal system with similar neuropathological changes has been described. We now report that the injection of kainic acid into the rat striatum causes neuronal degeneration, neurochemical alterations and behavioural responses resembling Huntington's chorea. This procedure could provide an animal model for the study of the disease.

Systemic administration of monosodium glutamate to immature mice causes degeneration of certain neurones of retina[6] and brain[7]. In brain, neurones with cell bodies in the arcuate nucleus, which selectively accumulates glutamate[8], degenerate whereas axons passing through or terminating in the region are unaffected[9]. The neurotoxic effects of glutamate may be due to its ability to depolarise neurones, for the neurotoxicity of several dicarboxylic and sulphur-containing amino acids structurally related to glutamate correlates with their neuroexcitatory action[10]. In addition, neuronal sensitivity to depolarisation by glutamate is limited to dendrites and soma while axons are unresponsive[11]. Systemic administration of kainic acid, a rigid analogue of glutamate with potent neuroexcitatory effects[12], causes degeneration of central nervous system (CNS) neurones in adult mice; but the attendant high incidence of seizures and death has limited its experimental use[13]. Intracerebral application of excitatory amino acids, however, induces localised neuronal degeneration with neuropathological characteristics similar to that resulting from peripheral administration[14,15].

As the sensitivity of neurones to the toxic effects of neuroexcitatory amino acids may be limited to soma and dendrites, it seemed that an intracerebral injection of kainic acid might cause degeneration of neurones with cell bodies near the injection site, but spare axons terminating in or passing through the area. We have tested the striatum because of its defined neuroanatomy and neurochemistry; it receives a dense innervation from dopaminergic neurones with cell bodies localised in the substantia nigra whereas the striatum itself is populated primarily by small neurones of which cholinergic and GABAergic (gamma aminobutyric acid) neurones are a major component[16]. Accordingly, we have examined the effects of a stereotaxic infusion of kainic acid into the striatum on the neurochemical markers for the intrinsic cholinergic and GABAergic neurones and the dopaminergic terminals in the region.

After anaesthesia with Equi-Thesin (0.6 ml; Jenson–Salisbury Labs), Sprague–Dawley rats (160 g) were positioned in a David Kopf small animal stereotaxic apparatus; and a 0.3-mm Hamilton cannula was inserted into the striatum through a burr hole in the calvarium (coordinates: 7.9 A; 2.6 L; 4.8 V). Drugs dissolved in 1 μl of artificial cerebrospinal fluid (CSF) (titrated to pH 7.4) were infused over 1 min; the cannula was then removed and the scalp
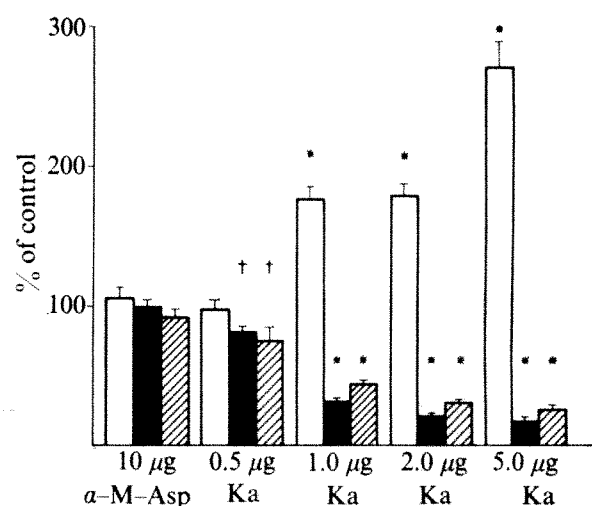
Fig. 1 The activities of tyrosine hydroxylase[23] (open bar), choline acetyltransferase[24] (solid bar) and glutamic acid decarboxylase[25] (striped bar) in the striatum were assayed 48 h after unilateral injection of various amounts of kainic acid (Ka) or α-methylaspartate (α-M-Asp). Results are presented in terms of percentage of the specific activity of the contralateral striatum with s.e.m. indicated. $N = 5$. $*P < 0.001$. $†P < 0.05$.

was apposed with sutures. At various times after treatment, the rats were killed; and the striata were dissected according to the method of Glowinski and Iversen[17].

In initial experiments, the effects of infusion of various amounts of kainic acid (0.1–10 μg) on the neurotransmitter-synthesising enzymes in the striatum were examined 48 h after injection (Fig. 1). With the injection of 0.5 μg, there was a 20% reduction in the activities of glutamic acid decarboxylase (GAD) and choline acetyltransferase (CAT). With increasing doses, there was a progressive reduction in the activities of GAD and CAT whereas the activity of tyrosine hydroxylase (TH) exhibited a dose-dependent increase. Two days after injection of 2 μg of kainic acid, the activities of GAD and CAT were reduced 80% whereas the activity of TH was increased 80%. Doses in excess of 5 μg of kainic acid did not further alter enzyme activities in the striatum but were associated with a high rate of mortality. These changes in enzyme activity were not caused by the injection procedure, for the infusion of 10 μg of α-methyl-aspartate, an analogue of glutamate without neuroexcitatory activity[18], and did not alter significantly the activities of these enzymes. In contrast to the effects of kainic acid, the infusion of copper sulphate (25 μg), a nonspecific toxic agent[19], caused a similar 60–65% reduction in the activities of TH as well as GAD and CAT.

It is possible that the 80% decrease in the activities of CAT and GAD and 80% increase in the activity of TH after injection of 2 μg of kainic acid reflected merely a transient neurochemical adaptation to excessive stimulation of glutamate receptors. To obtain further information, we examined several neurochemical markers for the cholinergic, GABAergic and dopaminergic neurones 10 d after infusion of 2.5 μg of kainic acid (Table 1). The injection elicited a 60–70% decrement in the activity of GAD, the levels of endogenous GABA, the activity of the synaptosomal high affinity uptake process for GABA. Similarly, the activity of CAT, the levels of endogenous acetylcholine and the activity of the synaptosomal uptake process for choline were reduced 60–70%. In contrast, the activity of TH was 50% above control; the levels of endogenous dopamine and the activity of the synaptosomal uptake process for dopamine were not significantly affected by the injection. In addition, the wet weight of the striatum and protein content were not changed by the treatment.

Table 1 Effects of kainic acid on neurochemical parameters of the striatum 10 d after injection

		Injected	Control	Δ
GABAergic neurones				
Glutamic acid decarboxylase	nmol mg^{-1} h^{-1}	4.13±0.8	12.9±0.9	−68*
GABA	nmol mg^{-1}	0.72±0.15	2.16±0.17	−67*
GABA uptake	pmol per mg per 4 min	7.1±1.2	16.5±1.3	−57*
Cholinergic neurones				
Choline acetyltransferase	nmol mg^{-1} h^{-1}	6.0±1	20±1	−70*
Acetylcholine	pmol mg^{-1}	20±4	68±4	−71*
Choline uptake	pmol per mg per 4 min	0.05±0.01	0.12±0.01	−60*
Dopaminergic neurones				
Tyrosine hydroxylase	pmol mg^{-1} hr^{-1}	340±25	232±21	+47*
Dopamine	pmol mg^{-1}	52±4	60±4	−13
Dopamine uptake	pmol per mg per 4 min	3.1±0.3	2.4±0.3	+29
Wet weight	mg	48±2	49±2	−2
Protein	µg per mg tissue	165±9	173±7	−5

Kainic acid (2.5 µg) was injected into the striatum 10 d before killing the rats. The activities of tyrosine hydroxylase[23], choline acetyltransferase[24], and glutamic acid decarboxylase[25] and the levels of GABA[26], acetylcholine[27], dopamine[28] and protein[29] were assayed. The uptake of labelled GABA[30], choline[31], and dopamine[32] was measured in washed P$_2$ fractions prepared from striatum. The contralateral uninjected striatum served as control $N = 6-12$.
*$P < 0.001$.

Some of the GABAergic neurones with cell bodies in the striatum provide an inhibitory pathway innervating the dopaminergic cell bodies in the substantia nigra[16,20]. In rats receiving a unilateral infusion of 2.5 µg of kainic acid in the striatum 1 week before death, the levels of endogenous GABA were reduced 67±3% ($N=5$; $P<0.001$) in the ipsilateral substantia nigra. In contrast, the activity of tyrosine hydroxylase was not significantly altered in the substantia nigra on the injected side. Thus, the striatonigral GABAergic pathway seems to be affected selectively by the striatal injection of kainic acid.

Although these marked neurochemical alterations 10 d after injection of kainic acid suggest specific degeneration of cholinergic and GABAergic neurones intrinsic to the striatum, histological evidence is required. Therefore histological sections stained with cresyl violet were prepared from whole forebrain at the level of the striatum of animals

Fig. 2 Micrographs of rat striatum showing the loss of neurones on the side injected with 2.5 µg of kainic acid (*b*) compared with the contralateral side (*a*) as control. Ten days after injection, the animals were killed by aortic perfusion with buffered formaldehyde. Sections (30 µm) were stained with cresyl violet. Neuronal perikarya are markedly reduced in number on the injected side which also shows marked gliosis. The lighter areas, which are bundles of internal capsule fibres, are unaffected by the treatment. The bar represents 100 µm.

injected unilaterally with 2.5 µg of kainic acid 10 d before death (Fig. 2). Around the site of injection there was generalised loss of approximately 95% of neuronal perikarya in a radius of 1.5 mm, with a gradation in loss with increasing distance. Thus, the lesion involved nearly the entire head of the caudate putamen but did not affect the tail of the caudate or the globus pallidus (our work in preparation). Although the needle track was visible, there was no evidence of inflammatory changes or necrosis at the injection sites, but there was a generalised increase in the number and prominence of glia. The overall structure of the striatum containing the bundles of internal capsule fibres seemed to be unchanged except for the loss of neurones.

The striking behaviour of animals treated with a unilateral injection of kainic acid in the striatum demonstrates the physiological significance of the lesion. For up to 24 h after injection, the rats exhibited marked rotatory behaviour away from the side of the injection, which would be consistent with the neuroexcitatory effects of the drug in the striatum[15]. Two days after injection, however, they reversed the direction of rotation towards the side of the injection. The ipsilateral rotatory behaviour, which persisted up to 10 d after injection, is similar to that produced by lesions in the dopaminergic nigrostriatal pathway[21]. Thus, the lesion caused by kainic acid in striatal neurones, which are presumably postsynaptic to the dopaminergic terminals[16,22], results in a behaviour similar to that induced by ablation of the dopaminergic innervation to the striatum.

The striatal lesions produced by kainic acid have particular relevance to Huntington's chorea because the major site of neuropathological change in this disorder is the striatum, where a marked loss of intrinsic neurones and gliosis occurs[1]. There are considerable decreases in the neurochemical markers for the cholinergic and GABAergic neurones in the striatum whereas the dopaminergic innervation to the region is relatively unaffected; in addition, there is a selective reduction in the activity of GAD and levels of GABA in the substantia nigra[2-5]. The neurochemical, histological and behavioural changes resulting from striatal infusion with kainic acid are remarkably similar to those observed in Huntington's chorea. Thus, lesions of the rat striatum produced with kainic acid may provide a useful animal model for Huntington's chorea, in which the biochemical, electrophysiological and pharmacological sequelae of the loss of striatal cholinergic and GABAergic neurones can be examined.

We thank Robert Zaczek for technical assistance and Dr Michael Kuhar and Naomi Taylor for doing the histological

studies. This work was supported by a grant to J.T.C. from USPHS.

JOSEPH T. COYLE
ROBERT SCHWARCZ

Departments of Pharmacology and Experimental
  Therapeutics and Psychiatry and the Behavioural Sciences,
Johns Hopkins University School of Medicine,
Baltimore, Maryland 21205

1 Bruyn, G. W., in *Handbook of Clinical Neurology, Diseases of Basal Ganglia*, 6 (edit. by Vinken, P. J., and Bruyn, G. W.), 298–377 (North-Holland, Amsterdam, 1968).
2 Perry, T. L., Hansen, S., and Kloster, M., *New Engl. J. Med.*, **288**, 337–342 (1973).
3 Bird, E. D., and Iversen, L. L., *Brain*, **97**, 457–472 (1974).
4 McGeer, P. L., and McGeer, E. G., *J. Neurochem.*, **26**, 65–76 (1976).
5 Stahl, W. L., and Swanson, P. D., *Neurology*, **24**, 813–819 (1974).
6 Lucas, D. R., and Newhouse, J. P., *Am. med. Ass. Archs Opthal.*, **58**, 193–204 (1957).
7 Olney, J. W., *Science*, **164**, 719–721 (1969).
8 Perez, V. J., and Olney, J. W., *J. Neurochem.*, **19**, 1777–1782 (1972).
9 Olney, J. W., Schainker, B., and Rhee, V., in *Hormones, Behavior and Psychopathology* (edit. by Sachar, E.), 153–158 (Raven, New York, 1976).
10 Olney, J. W., Rhee, V., and Ho, O. L., *Expl Brain Res.*, **14**, 61–76 (1972).
11 Zieglgansberger, W., and Puil, E. A., *Expl Brain Res.*, **17**, 35–49 (1973).
12 Buu, N. T., and Van Gelder, N. M., *Gen. Pharmac.*, **7**, 5–14 (1976).
13 Olney, J. W., Rhee, V., and Ho, O. L., *Brain Res.*, **77**, 507–512 (1974).
14 Van Harreveld, A., and Fifova, E., *Expl molec. Path.*, **15**, 61–81 (1971).
15 Olney, J. W., Sharpe, L. G., and de Gubareff, T., *Neurosci. Abstr.*, **1**, 371 (1975).
16 McGeer, P. L., and McGeer, E. G., in *GABA in Nervous System Function* (edit. by Roberts, E., Chase, T. N., and Tower, D. B.), 487–496 (Raven, New York, 1976).
17 Glowinski, J., and Iversen, L. L., *J. Neurochem.*, **13**, 655–669 (1966).
18 McLennan, H., in *Handbook of Psychopharmacology: Amino Acid Neurotransmitters*, **4** (edit. by Iversen, L. L., Iversen, S. D., and Snyder, S. H.), 211–228 (Plenum, New York, 1975).
19 Javoy, F., Sotelo, C., Herbert, A., and Agid, Y., *Brain Res.*, **102**, 201–215 (1976).
20 Fonnum, F., Grofora, I., Rinvik, E., Storm-Mathiesen, J., and Walburg, F., *Brain Res.*, **71**, 77–92 (1974).
21 Ungerstedt, U., *Acta physiol. scand.*, Suppl., **367**, 1–22 (1971).
22 Racagni, G., Cheney, D. L., Trabucchi, M., and Costa, E., *J. Pharmac. exp. Ther.*, **196**, 323–332 (1976).
23 Coyle, J. T., *Biochem. Pharmac.*, **21**, 1935–1944 (1972).
24 Bull, G., and Oderfeld-Nowak, B., *J. Neurochem.*, **19**, 935–947 (1971).
25 Wilson, S. H., *et al.*, *J. biol. Chem.*, **247**, 3159–3169 (1972).
26 Graham, L. T., and Aprison, M. H., *Analyt. Biochem.*, **15**, 487–497 (1966).
27 Goldberg, A. M., and McCamman, R. E., *J. Neurochem.*, **20**, 1–8 (1973).
28 Coyle, J. T., and Henry, D., *J. Neurochem.*, **21**, 61–67 (1973).
29 Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, P. J., *J. biol. Chem.*, **193**, 265–275 (1975).
30 Coyle, J. T., and Enna, S. J., *Brain Res.* **111**, 119–133 (1976).
31 Simon, J. R., Atweh, S., and Kuhar, M. J., *J. Neurochem.*, **26**, 909–922 (1976).
32 Holz, R. W., and Coyle, J. T., *Molec. Pharmac.*, **10**, 746–758 (1974).

Evidence for a role of central serotonergic neurones in digitalis-induced cardiac arrhythmias

EVIDENCE indicates that digitalis drugs administered intravenously increase central sympathetic outflow and that this in turn results in cardiac arrhythmias[1–6]. The central nervous system transmitter(s) that mediates this effect is not known. According to Saito *et al.*, noradrenaline, the primary transmitter studied so far, is not involved in the case of guinea pigs[7]. 5-Hydroxytryptamine (5-HT) however, may be an important chemical mediator regulating central sympathetic outflow[8,9], and it is involved in the respiratory arrest induced in rats by intravenous administration of digitoxigenin[10]. We now report evidence that suggests that brain 5-HT is involved in digitalis-induced cardiac arrhythmias.

Cats (1.6–3.6 kg) were anaesthetised with alpha chloralose (70–80 mg kg^{-1} intravenously) and artificially ventilated with room air. The femoral artery and vein were catheterised for recording blood pressure and administering drugs, respectively. Rectal temperature was maintained between 36° and 38 °C by warming the cat with radiant heat. The electrocardiogram and the arterial blood pressure were recorded. The digitalis preparation, deslanoside, was administered by a continuous intravenous infusion of 2 µg kg^{-1} min^{-1}. Doses necessary to produce the following endpoints were obtained: ventricular arrhythmia, ventricular tachycardia and ventricular fibrillation. The role of brain 5-HT in the arrhythmogenic effects of deslanoside was assessed by determining: (1) the effect of pretreatment with *p*-chlorophenylalanine (PCPA) on the doses of deslanoside

needed to produce ventricular arrhythmias (PCPA was administered as 300 mg kg^{-1} intraperitoneally each day for 3 d before the animals were intoxicated with deslanoside); (2) the effect of 5,7-dihydroxytryptamine (5,7-DHT) on the doses of deslanoside needed to produce ventricular arrhythmias (5,7-DHT was administered as a single dose of either 100 µg kg^{-1} to three cats, or 200 µg kg^{-1} to two cats, into the anterior horn of the left lateral ventricle 8 d before intoxication with deslanoside), and (3) the effect of methysergide as both pretreatment (3 mg kg^{-1} intravenously, 10 min before deslanoside infusion and repeated 45 and 85 min after deslanoside infusion started) and as a treatment (3 mg kg^{-1} intravenously, 1–6 min after a ventricular arrhythmia had been evoked by deslanoside).

To assess the extent and selectivity of the depletion of brain 5-HT by PCPA and 5,7-DHT, the brains of cats intoxicated with deslanoside were removed immediately after the onset of ventricular fibrillation. The brains were placed on ice and the medulla–pons, hypothalamus and superior and inferior colliculi were dissected out, frozen on dry ice and stored at −20 °C until assayed (up to 3 weeks). Biogenic amines, 5-HT and noradrenaline were assayed by a modification of the procedures of Chang *et al.*[11] and Maickel *et al.*[12], as described in detail by Zenker *et al.*[13]. Results are expressed as the mean µg of amine per g of tissue ±s.e.m.

Fig. 1 Effect of methysergide on deslanoside-induced ventricular arrhythmia. *a*, Electrocardiogram (ECG) and arterial blood pressure (BP) tracings before intoxication with deslanoside. *b*, Effects of a cumulative dose of deslanoside (200 µg kg^{-1}) on ECG and BP just before the development of a ventricular arrhythmia. *c*, Deslanoside-induced ventricular arrhythmia. *d*, ECG and BP tracings 2.5 min after 3 methysergide (3 mg kg^{-1}) was administered.
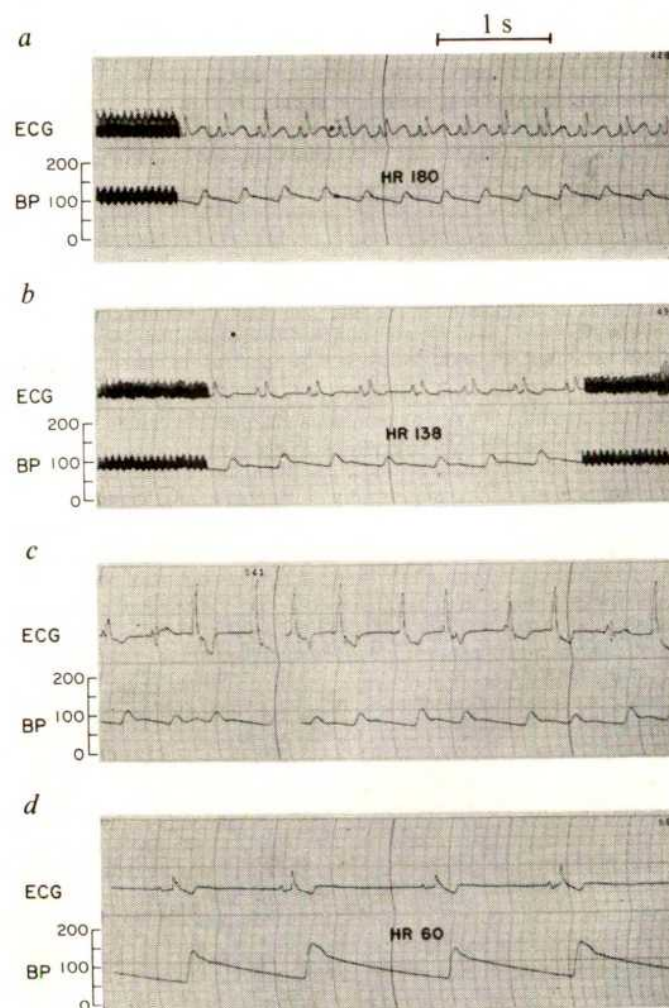
Table 1 Effect of p-chlorophenylalanine, 5,7-DHT and methysergide pretreatment on the capacity of deslanoside to produce ventricular arrhythmias

| Group | No. in each group | Heart rate and blood pressure before deslanoside infusion | | Doses of deslanoside to produce cardiac arrhythmias ($\mu g\ kg^{-1}$) | | |
		Heart rate (beats per min)	Mean blood pressure (mmHg)	Ventricular arrhythmia	Ventricular tachycardia	Ventricular fibrillation
Control	21	193 ± 5.8	136 ± 7.3	165 ± 8.5	178 ± 8.5	212 ± 11.3
p-Chlorophenylalanine pretreatment	7	186 ± 9.8	132 ± 7.8	$238 \pm 22.8^*$	$246 \pm 22.7^*$	$277 \pm 18.6^*$
5,7-Dihydroxytryptamine pretreatment	5	190 ± 6.4	123 ± 16.4	$219 \pm 10.6^*$	$223 \pm 10.4^*$	258 ± 15.1
Methysergide pretreatment	5	$131 \pm 12.3^*$	$96 \pm 7.7^*$	$232 \pm 29.0^*$	$249 \pm 27.5^*$	$317 \pm 35.4^*$

Controls were 12 animals given deslanoside only, three given saline intraperitoneally for 3 d and then intoxicated with deslanoside (to control for p-chlorophenylalanine injections), and six given saline into the anterior horn of the left lateral ventricle and 8 d later intoxicated with deslanoside (to control for the 5,7-DHT injection). Values between the three control groups were not significantly different so the data were pooled.
$^*P < 0.05$ when data are compared with corresponding control data.

of triplicate determinations. Using a Farrand Fluorescent MK-1 spectrophotofluorometer, the fluorescent intensity of the biogenic amines was determined in 1-cm quartz cells.

The pretreatment studies (Table 1) showed that animals given any one of the three agents that interfere with brain 5-HT function required a significantly larger dose of deslanoside to produce ventricular arrhythmias. PCPA and 5,7-DHT inhibit tryptophan hydroxylase and destroy central serotonergic nerves, respectively; animals pretreated with these drugs exhibited a significant decrease in the 5-HT content in all brain regions analysed (Table 2). Depletion

Taken together, these results suggest that central serotonergic neural systems are involved in the central arrhythmogenic effect of digitalis. Central serotonergic neural systems have been demonstrated by some investigators to have an important influence on central sympathetic outflow[8,9,19]. For example, Wing and Chalmers[8] reported that intracisternal administration of 5,6-DHT prevents the increases in arterial blood pressure and heart rate produced by sinoaortic denervation in unanaesthetised rabbits. This suggests that activity in central serotonergic nerves increases central sympathetic outflow. Wing and Chalmers did not

Table 2 Effect of p-chlorophenylalanine and 5,7-DHT on 5-HT and noradrenaline content of selected brain regions

| Group | No. in each group | Brain region | | | | | |
| | | Medulla–pons | | Hypothalamus | | Colliculi | |
		5-HT ($\mu g\ g^{-1}$)	Noradrenaline ($\mu g\ g^{-1}$)	5-HT ($\mu g\ g^{-1}$)	Noradrenaline ($\mu g\ g^{-1}$)	5-HT ($\mu g\ g^{-1}$)	Noradrenaline ($\mu g\ g^{-1}$)
Controls (no drug)	6	0.73 ± 0.03	0.91 ± 0.05	1.53 ± 0.04	1.30 ± 0.04	0.87 ± 0.03	0.28 ± 0.02
Animals intoxicated with deslanoside	6	0.75 ± 0.03	0.85 ± 0.04	1.53 ± 0.04	1.20 ± 0.04	0.86 ± 0.03	0.31 ± 0.03
Animals pretreated with p-chlorophenylalanine and intoxicated with deslanoside	3	$0.28 \pm 0.04^*$	0.87 ± 0.03	$0.22 \pm 0.02^*$	1.32 ± 0.04	$0.06 \pm 0.02^*$	0.29 ± 0.03
Animals pretreated with 5,7-DHT and intoxicated with deslanoside	5	$0.51 \pm 0.05^*$	0.86 ± 0.05	$0.83 \pm 0.16^*$	1.39 ± 0.05	$0.43 \pm 0.11^*$	0.29 ± 0.03

$^*P < 0.05$ when data are compared with either the control group or the animals intoxicated with deslanoside.

of 5-HT content with these drugs seemed to be selective, as no significant changes in noradrenaline were observed.

Deslanoside had no significant effect on brain 5-HT levels (Table 2). This was not unexpected, for treatments that enhance central serotonergic neural activity do not alter 5-HT levels although they increase turnover[14].

We tested methysergide for antiarrhythmic effects because this drug rapidly antagonises 5-HT receptors after intravenous administration[15]. It was administered intravenously (as a bolus of 3 mg kg⁻¹) to three control cats with a sustained ventricular tachycardia induced by deslanoside. In each cat, methysergide converted the existing ventricular arrhythmias to sinus rhythm (Fig. 1).

Pretreatment with PCPA and 5,7-DHT had no significant effect on baseline heart rate and arterial pressure. Pretreatment with methysergide resulted in a significant reduction in both heart rate and arterial pressure (Table 1).

We have also found[16] that doses of deslanoside that produce lethal ventricular arrhythmias significantly increase 5-hyrdoxyindoleacetic acid tissue concentration and tryptophan hydroxylase activity in the hypothalamus, amygdala and colliculi. This indicates that deslanoside produces biochemical changes in the central serotonergic system that are consistent with the drug activating serotonergic neurones[17]. We have also found[18] that electrical stimulation of midbrain raphe areas causes similar kinds of arrhythmias and biochemical changes (changes in 5-hydroxyindoleacetic acid and tryptophan hydroxylase activity), as does deslanoside.

observe drastic changes in baseline heart rate or arterial pressure with 5,6-DHT, indicating that normal sympathetic outflow may not be greatly affected by depleting brain 5-HT. In the study reported here, no significant changes in these baseline cardiovascular parameters were observed with either 5,7-DHT or PCPA. Thus interference with central 5-HT synthesis or selective degeneration of central 5-HT nerve endings may prevent noxious influences (digitalis overdose or sinoaortic denervation) from deranging central sympathetic outflow without disrupting normal tonic central sympathetic activity.

CINDA J. HELKE
JANETTE DIAS SOUZA
BETTY L. HAMILTON
VICTOR H. MORGENROTH, III
RICHARD A. GILLIS

Departments of Pharmacology and Anatomy,
Georgetown University,
Schools of Medicine and Dentistry,
Washington, DC 20007

1 Gillis, R. A., *Science*, 166, 508 (1969).
2 McLain, P. L., *Int. J. Neuropharmac.*, 8, 379 (1969).
3 Gillis, R. A., McClellan, J. R., Sauer, T. S., and Standaert, F. G., *J. Pharmac. exp. Ther.*, 179, 599 (1971).
4 Gillis, R. A., Raines, A., Sohn, Y. J., Levitt, B., and Standaert, F. G., *J. Pharmac. exp. Ther.*, 183, 154 (1972).

5 Garvey, H. L., in *Recent Advances in Studies on Cardiac Structure and Metabolism*, 1 (edit. by Bajusz, E., and Rona, G.), 563 (University Park Press, London, 1972).
6 Evans, D. E., and Gillis, R. A., *J. Pharmac. exp. Ther.*, **195**, 577 (1975).
7 Saito, H., Otani, T., Shudo, I., and Tanabe, T., *Jap. J. Pharmac.*, **24**, 923 (1974).
8 Chalmers, J. P., *Circulation Res.*, **36**, 469 (1975).
9 Wing, L. M. H., and Chalmers, J. P., *Circulation Res.*, **35**, 504 (1974).
10 Buterbaugh, G. G., and Spratt, J. L., *J. Pharmac. exp. Ther.*, **175**, 121 (1970).
11 Chang, C. C., *Int. J. Neuropharmac.*, **3**, 634 (1964).
12 Maickel, R. H., Cox, J., Saillant, J., and Miller, F. P., *Int. J. Neuropharmac.*, **7**, 275 (1968).
13 Zenker, N., Morgenroth III, V. H., Wright, J., King, J. W., and Arnett, C. D., *Biochem. Pharmac.*, **25**, 585 (1976).
14 Shields, P. J., and Eccleston, D., *J. Neurochem.*, **19**, 265 (1972).
15 Clineschmidt, B. V., and Anderson, E. B., *Expl Brain Res.*, **11**, 157 (1970).
16 Morgenroth, V. H. III, Helke, C. J., Dias Souza, J., Hamilton, B. L., and Gillis, R. A., *Neurosci. Abst.* (in the press).
17 Sheard, M. H., and Aghajanian, G. E., *J. Pharmac. exp. Ther.*, **163**, 425 (1968).
18 Gillis, R. A., Helke, C. J., Hamilton, B. L., and Morgenroth, V. H., III, *Neurosci. Abstr.* (in the press).
19 Lambert, G., Friedman, E., and Gershon, S., *Life Sci.*, **17**, 915 (1975).

Absolute sensitivity of rod bipolar cells in a dark-adapted retina

THE visual system of vertebrates mediated by rods is capable of detecting a few quanta of light[1,2]. The work reported here is concerned with the voltage response to flashes of light in individual bipolar cells, to which the rods send their signals and which, in turn, signal to amacrine and ganglion cells. We measured the absolute sensitivity of rod bipolar cells in the dogfish retina. The photoisomerisation of a single rhodopsin chromophore in only one rod out of 100 leads to a response of 1 mV in the bipolar cell.

Intracellular recordings were made in the eyecup preparation of the dogfish, *Scyliorhinus canicula*, dissected in dim red light from animals dark-adapted for at least 3 h. This species has visual receptors which are almost exclusively rods, containing a photopigment[3] with maximum absorbance at 500 nm. The light source was a tungsten–iodine lamp operated from a stabilised d.c. supply and filtered by a bandpass interference filter with peak transmission at 495 nm. The irradiance of the source through the bandpass and heat-absorbing filters was calibrated with a radiometer (Tektronix J16/J6502 radiance probe), and checked by a photomultiplier in single-photon counting mode. Dark adaptation was maintained by using dim test flashes to locate bipolar cells. Cells were penetrated in the region of retina backed by the tapetum lucidum. The responses reported here were obtained with full-field illumination.

The identification of the cells as bipolar cells is based on criteria of depth, waveform and receptive field. The cells were

Fig. 1 Responses of a bipolar cell in the dark-adapted retina of the dogfish *Scyliorhinus canicula* to 18-ms flashes of light of varying intensity. An increase of internal potential (depolarisation) is indicated by an upward deflection. The timing of the flash is shown beneath the superimposed traces. The numbers near each trace represent the number of photons μm^{-2} at 495 nm delivered by the flash. The internal potential of the cell in the dark was -40 mV.
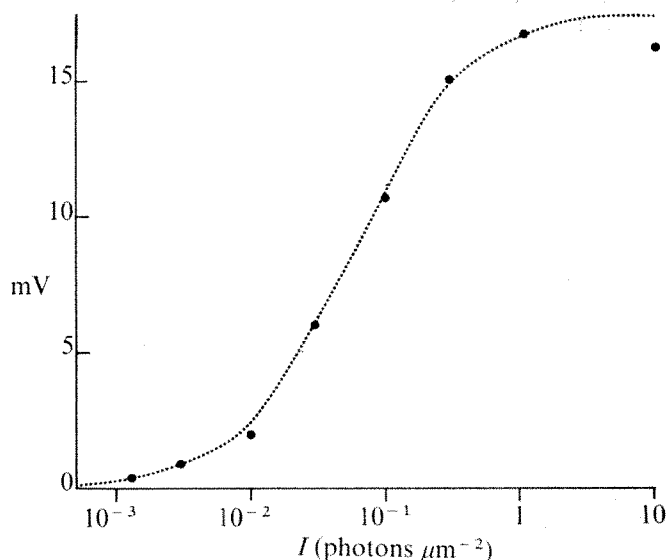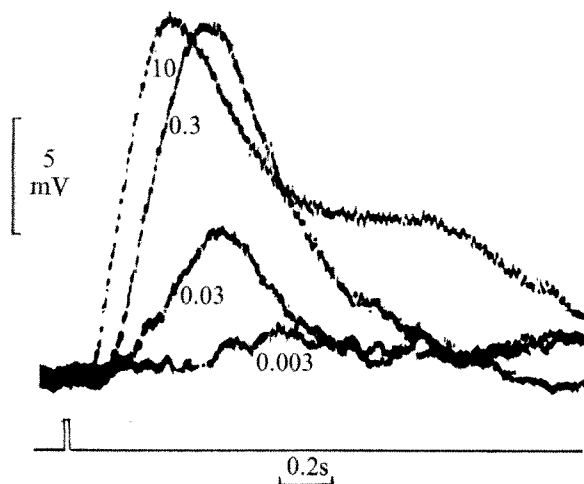
Fig. 2 The relationship between light intensity and peak of the response for the bipolar cell shown in Fig. 1. Note that the scale for light intensity is logarithmic. The three points shown at the lowest light levels were obtained by signal averaging, each point being the average of eight responses.

impaled at depths of 70–100 μm from the vitreal surface of the retina, often in close proximity to a horizontal cell. The response to light flashes (Fig. 1) consists of a depolarisation graded with the light intensity, with a maximum response of up to 30 mV. The response characterised by a phasic peak, which declines to a plateau, is evoked at higher light intensities. This is the response described by Kaneko for bipolar cells, identified by staining with Procion dyes in the retina of the related species, *Mustelis canis*[4]. As in bipolar cells in other retinae, there is an 'off' response which follows the cessation of a long, near-saturating pulse of light[5]. The receptive field for the centre response of these cells is less than 500 μm and in some cells less than 200 μm. An antagonistic effect produced by illumination of the surround is often used to identify bipolar cells; the cells reported here, however, have only a weak centre-surround organisation. There are reasons to suppose that dark adaptation diminishes the effect of the surround[6].

To analyse the sensitivity of the cells, the responses to brief flashes of light were recorded at low light levels in the region over which the response was proportional to intensity, as well as at higher light levels. The small signals evoked by dim lights were averaged by means of a computer. Figure 2 shows the intensity–response curve for the cell whose responses were illustrated in Fig. 1. The line of Fig. 2 fitted to the experimental points is of the form

$$V = V_{max}\frac{I}{I+I_{\frac{1}{2}}} \qquad (1)$$

a hyperbolic relationship between the response, V, and the light intensity, I. For this cell $V_{max} = 17.5$ mV and the half-saturating intensity, $I_{1/2} = 0.06$ photons μm^{-2} per flash. The absolute sensitivity of the cell[7] was determined as the ratio of response to intensity of light, within the range over which the response was proportional to light intensity. The flash sensitivity for this cell was 310 mV photon^{-1} μm^2. The results for other bipolar cells are given in Table 1. There was considerable variation in sensitivity and $I_{1/2}$ values among cells, but the flash sensitivity consistently exceeded 100 mV photon^{-1} μm^2, provided that V_{max} was greater than 8 mV.

To convert these figures to the voltage response that would be produced by the photoisomerisation of a single rhodopsin

Table 1 Flash sensitivity of rod bipolar cells and the semi-saturation value of light intensity, $I_{1/2}$

V_{max} (mV)	Flash sensitivity (mV photon^{-1} μm^2)	(mV rod per isomerisation)	$I_{1/2}$ (isomerisation per rod)
17.5	310	103	0.18
23.8	330	110	0.42
21.8	200	67	0.33
14.1	373	124	0.55
13.8	137	46	1.20
25.5	170	57	0.63
8.1	476	159	0.18
9.3	228	76	0.55
18.5	323	108	0.19
Mean	283	95	0.47

For some cells, the slope of the intensity–response curve near $I_{1/2}$ was significantly less steep than would be indicated by equation (1). The extent of the deviation may be gauged by comparing the flash sensitivity derived from the response to dim flashes with the value $V_{max}/I_{1/2}$. The internal potential of the cells in the dark ranged between -40 and -62 mV.

chromophore in each rod, an estimate of the photoisomerisation cross section (also called the effective collecting area[7]) is required. The geometric cross section of *Scyliorhinus* rod outer segments suspended in vitreal fluid is 7.1 μm^2 and their length is 28 μm. On the assumption of a specific axial density[8] for rhodopsin of 0.014 μm^{-1}, a quantum efficiency for photoisomerisation of 0.67 (ref. 9), a reflectivity of the tapetum of 90% (ref. 3) and the further assumption that all the incident light passes into the outer segments, the photoisomerisation cross section would be 3.7 μm^2. This figure takes no account of the fact that some of the light will pass between photoreceptors as well as being scattered by other retinal layers. We therefore tentatively assign a value of 3 μm^2 for the photoisomerisation cross section. This value has been used to derive the flash sensitivities of the rod bipolar cells (given as the response in mV per photoisomerisation in each rod) and to estimate the number of photoisomerisations per rod necessary for half saturation of the response of a bipolar cell. On average, the response is half-saturated when one rhodopsin chromophore is isomerised in just one out of two rods and there is a response of ~ 1 mV in the bipolar cell when there has been one photoisomerisation in one out of 100 rods (Table 1).

Because of their small size, it has not been possible to determine the flash sensitivity of dogfish rods. The highest values reported for flash sensitivities of uniformly illuminated rods were obtained by Fain[10] in *Bufo marinus* (0.68 mV per photoisomerisation) and by Copenhagen and Owen[11] in the snapping turtle (0.72 mV per photoisomerisation). There is no reason to suppose that the rods of dogfish are unusually sensitive. Measurement of the externally recorded receptor potential of the electroretinogram yields a value for the number of photoisomerisations for half saturation similar to that reported for other vertebrate rods[10-14]. It is possible that the isomerisation of a chromophore in a dogfish rod may lead to a somewhat larger voltage change than has been observed in rods of larger diameter. If photoisomerisation produces the same underlying membrane conductance change in all rods, the voltage change produced would then depend on the dimensions of the cell, varying approximately as 1/diameter if the rod were regarded as a short cable. On this basis, the flash sensitivity for dogfish rods might be about 2 mV per photoisomerisation. If rods are electrically coupled[10,11,15], then at low light levels when only a few rods absorb a photon, the voltage response in any one rod will be only a small fraction of what would be obtained in an electrically independent rod.

These measurements allow some deductions about signal transmission between rods and bipolar cells to be made. The dynamic voltage gain in signal transmission is the ratio of the flash sensitivity of the bipolar cell to that of the rods; in this case, yielding a gain of about 50. Schwartz[16] also concluded that there was a large increase in signal amplitude in transmission from cones to bipolar cells in the turtle retina. A dynamic voltage gain of 50 can be compared with a value of

about 3, applicable to a number of conventional chemical synapses[17]. In both cases, it must be assumed that the postsynaptic cell linearly sums the transmitter-modulated conductance increments contributed by all the presynaptic cells which converge on it. The maximum dynamic gain can be specified in terms of the fractional contribution of transmitter-modulated conductance to the total conductance of the postsynaptic cell (and, therefore, does not depend explicitly on convergence). The higher dynamic gain at the rod–bipolar cell synapse remains to be explained.

This work was supported by the MRC.

J. F. ASHMORE
GERTRUDE FALK

Department of Biophysics,
University College London,
Gower Street, London WC1E 6BT, UK

1 Hecht, S., Shlaer, S., and Pirenne, M. H., *J. gen. Physiol.*, **25**, 819–840 (1942).
2 Barlow, H. B., Levick, W. R., and Yoon, M., *Vision Res. Suppl.*, **3**, 87–101 (1971).
3 Denton, E. J., and Nicol, J. A. C., *J. mar. biol. Ass. U.K.*, **44**, 219–258 (1964).
4 Kaneko, A., *J. Physiol., Lond.*, **213**, 95–105 (1971).
5 Ashmore, J. F., and Falk, G., *J. Physiol., Lond.*, **258**, 39–40P (1976).
6 Barlow, H. B., Fitzhugh, R., and Kuffler, S. W., *J. Physiol., Lond.*, **137**, 338–354 (1957).
7 Baylor, D. A., and Hodgkin, A. L., *J. Physiol., Lond.*, **234**, 163–198 (1973).
8 Liebman, P. A., and Entine, G., *Vision Res.*, **8**, 761–775 (1968).
9 Dartnall, H. J. A., *Vision Res.*, **8**, 339–358 (1968).
10 Fain, G. L., *Science*, **187**, 838–841 (1975).
11 Copenhagen, D. R., and Owen, W. G., *J. Physiol., Lond.*, **259**, 251–282 (1976).
12 Penn, R. D., and Hagins, W. A., *Biophys. J.*, **12**, 1073–1094 (1972).
13 Fain, G. L., and Dowling, J. E., *Science*, **180**, 1178–1181 (1973).
14 Dowling, J. E., and Ripps, H., *J. gen. Physiol.*, **60**, 698–719 (1972).
15 Schwartz, E. A., *J. Physiol., Lond.*, **257**, 379–406 (1976).
16 Schwartz, E. A., *J. Physiol., Lond.*, **236**, 211–234 (1974).
17 Falk, G., and Fatt, P., in *Handbook of Sensory Physiology*, 7 (edit. by Dartnall, H. J. A.), 200–244 (Springer, Heidelberg, 1972); in *Lecture Notes in Biomathematics*, 4 (edit. by Conrad, M., Güttinger, W., and Dal Cin, M.), 171–197 (Springer, Heidelberg, 1974).

Identification of actin-binding protein in membrane of polymorphonuclear leukocytes

A NEW protein with a high molecular weight and found in rabbit alveolar macrophages[1], causes actin filaments in sucrose solutions to form a solid gel when warmed to room temperature. With transmission electron microscopy, the actin filaments were observed to be bound into branching interconnecting bundles[2]. The new protein was called actin-binding protein (ABP). Similar consistency changes, which are due to gelation of actin associated with other proteins, have been observed in sucrose extracts of *Acanthamoeba* and glycerol extracts of sea urchin eggs[3,4]. ABP was subsequently isolated from extracts of chronic myelogenous leukaemia (CML) leukocytes[5]. The CML leukocyte ABP had identical properties to the analogous protein isolated from the rabbit alveolar macrophage. The CML leukocyte ABP in the presence of actin produced a gel that structurally resembled the hyaline ectoplasm of a pseudopod. Since particle contact with the phagocytic membrane leads

to generation of a pseudopod, it seemed reasonable that ABP might be associated with the membrane. Using an antibody directed against CML leukocyte ABP coupled with indirect immunological stains, we were able to localise ABP to the cytoplasmic surface of the membrane of normal polymorphonuclear leukocytes.

We prepared[6] rabbit IgG against purified CML leukocyte ABP (Fig. 1). The antibody activity was eliminated by absorption with either purified ABP or intact normal polymorphonuclear (PMN) leukocytes. Furthermore, no cross reactivity was demonstrated in an Ochterlony plate with two different preparations of myosin antiserum. We had noted previously that ABP antibody inhibited the gelation of CML leukocyte extracts as well as of purified leukocyte actin filaments in the presence of ABP[5].

In rabbit alveolar macrophages, contact of an ingestible particle with the cell surface generates more extractable ABP[2]. These findings suggest that ABP might be associated with the plasma membrane of phagocytes. To verify whether human leukocyte ABP is found in PMN leukocyte membranes, we prepared phagolysosomes after allowing the cells to phagocytose opsonised lipopolysaccharide-coated paraffin oil droplets[7]. The phagolysosome forms by fusion of lysosomal granule membranes with inverted sections of the plasma membrane. These 'inside-out' membrane preparations were exposed to purified anti-ABP IgG or normal rabbit IgG and then treated with goat anti-rabbit IgG (γ chain) conjugated with fluorescein isothiocyanate (FITC; Hyland Laboratory, Costa Mesa, California). Diffuse membrane fluorescence was observed with the anti-ABP IgG (Fig. 2). Since the plasma membrane Fc receptors were hidden inside these phagolysosomes, there was no observable fluorescence with normal rabbit IgG. Staining was also eliminated from ABP antibodies by absorbing the antibody with purified ABP or intact PMN leukocytes.

The nonspecific fluorescence of intact cells prevented us from using the same technique to identify ABP on their plasma membranes. For this study, we used an immunoperoxidase stain[8] which eliminated the problem. Peripheral

Fig. 2 Phagocytic vesicles were incubated with ABP IgG for 20 min at 25 °C. The vesicles were washed three times in Krebs–Ringer phosphate, pH 7.4, and suspended in goat anti-rabbit IgG (γ chain) conjugated with fluorescein for 20 min; then they were washed again. The resulting phagocytic vesicle was observed after settling on to a glass coverslip using a Zeiss epiilluminator with a FITC interference filter, 450 dichroic mirror, 53 barrier filter, and a $\times 100$ objective.

Fig. 1 Immunodiffusion of human leukocyte ABP (centre well) against anti-ABP IgG (1) and against 1:2 dilution of the anti-ABP IgG (4); two different anti-leukocyte myosin antisera (5 and 6); anti-ABP antisera absorbed with intact PMN leukocytes (2); and anti-ABP antisera absorbed with purified ABP antigen (3).

blood leukocytes were treated with peroxide and methanol to block endogenous peroxidase. The cells were incubated with anti-ABP IgG, washed, and incubated with goat anti-rabbit IgG conjugated with horseradish peroxidase and washed again. The fixed dried smears were then stained for peroxidase. Those cells incubated with anti-ABP IgG had maximal reactivity at the cell margins (Fig. 3). No

Fig. 3 Peripheral blood cells stained with normal rabbit IgG (*a*) and with anti-ABP IgG (*b*). Note the positive reaction at the periphery of the cells.

a

b
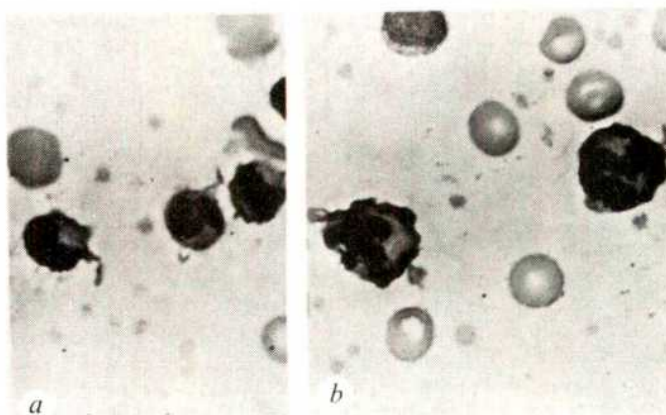
staining of PMN leukocytes occurred with a normal IgG nor of lymphocytes or monocytes with the specific antibody.

Actin, myosin and ABP have been identified by polyacrylamide gel electrophoresis in enriched rabbit alveolar membrane fractions[9]. Whether ABP is an integral protein or largely associated with surfaces of the leukocyte membrane is not ascertainable from our studies. It is, however, tempting to believe that ABP by its membrane location could transduce surface recognition of particles into mechanical events, leading to actin polymerisation and the generation of pseudopodia.

This study was supported by the NIH and the James Whitcomb Riley Memorial Association.

LAURENCE A. BOXER
SUSAN RICHARDSON

Department of Pediatrics,
James Whitcomb Riley Hospital for Children,

ALTON FLOYD

Department of Anatomy,
Indiana University School of Medicine,
1100 West Michigan Street,
Indianapolis, Indiana 46202

1 Hartwig, J. H., and Stossel, T. P., *J. biol. Chem.*, 250, 5696–5705 (1975).
2 Stossel, T. P., and Hartwig, J. H., *J. Cell Biol.*, 68, 602–619 (1976).
3 Pollard, T. K., *J. Cell Biol.*, 68, 579–601 (1976).
4 Kane, R. E., *J. Cell Biol.*, 66, 305–315 (1975).
5 Boxer, L. A., and Stossel, T. P., *J. clin. Invest.*, 56, 964–976 (1976).
6 Boxer, L. A., and Stossel, T. P., *J. clin. Invest.*, 53, 1534–1545 (1974).
7 Stossel, T. P., Pollard, T. D., Mason, R. J., and Vaughn, M., *J. clin. Invest.*, 50, 1745–1757 (1971).
8 Mason, D. T., Farrell, C., and Taylor, C. R., *Br. J. Haemat.*, 31, 361–370 (1975).
9 Pincus, S. H., and Stossel, T. P., *Clin. Res.*, 24, 109A (1976).

Transient kinetics of (Na+ +K+)-ATPase studied with a fluorescent substrate

WE describe experiments, with a purified (Na[+]+K[+])-ATPase, designed to investigate details of the enzymatic process that couples ATP hydrolysis to the active transport of Na[+] and K[+] ions. We have found that formycin triphosphate[1-3] (FTP), an analogue of ATP previously used in studies of myosin ATPase[4], is a substrate for the (Na[+]+K[+])-ATPase, and that the binding of FTP or of formycin diphosphate (FDP) to the enzyme is accompanied by a two- to threefold enhancement of nucleotide fluorescence. The change in strength of the fluorescent signal has allowed us to measure equilibrium binding of the nucleotides to the enzyme, and using stopped-flow fluorimetry we have been able to measure the rates of binding and release. FTP and FDP bind with a high affinity and may be displaced by excess ATP or by K[+] ions. During turnover, at least with FTP concentrations up to 24 μM, the FDP is released (or its fluorescence is quenched) before the rate-limiting step of the overall reaction, both in the presence and absence of K[+] ions. When turnover is prevented by the absence of magnesium, the enzyme is still able to change its conformation according to whether Na[+] or K[+] is the predominant alkali-metal ion[5]. Because the Na and K conformations have different nucleotide affinites, we have been able to measure the rates of the conformational changes by monitoring the decrease or increase in fluorescence associated with the net release or binding of nucleotide. Finally, the effects of alterations in FTP concentration on the time course of the fluorescence changes have provided a clue to the non-phosphorylating role of ATP.

All experiments were done at about 21 °C with a (Na[+]+K[+])-ATPase from pig kidney, purified by the simpler method of Jørgensen[6]. The enzyme had a specific activity of about 17 μmol
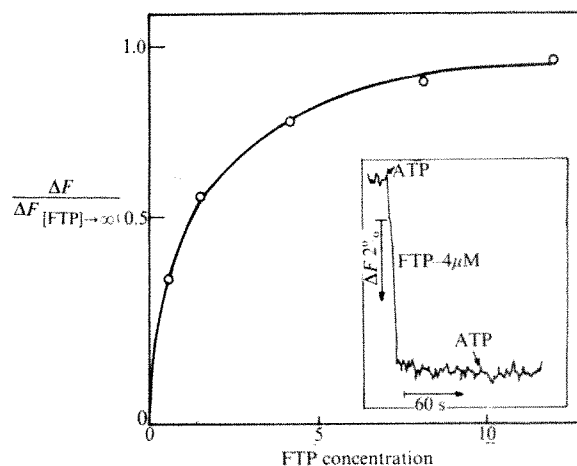
Fig. 1 Binding of FTP to (Na[+]+K[+])-ATPase as a function of FTP concentration. (Na[+]+K[+])-ATPase prepared from pig kidney by the simpler procedure of Jørgensen[6] was washed to remove traces of ATP and suspended at 60 μg ml[-1] in 2.5 ml of a solution containing NaCl 50 mM, EDTA (Tris salt, pH 7.7) 5 mM, and FTP (triethylamine salt) as indicated. ATP was added in 2 μl steps to give a final concentration of 25 μM. Fluorescence was measured at room temperature in a Farrand fluorimeter modified to allow additions to be made during continuous stirring and recording. Fluorescence was excited at 310 nm, and the light emitted at 380 nm was measured using 10-nm slits. A Wratten 18A stray-light filter was placed after the excitation monochromator, and a Schott W.G.365 stray-light filter before the emission monochromator. With 4 μM FTP, about half of the emitted light came from intrinsic protein fluorescence and scattering. Inset: reduction in fluorescence (ΔF) produced by the addition of ATP. Note that a second addition caused no further change, and that with 4 μM FTP, ΔF was only about 4% of total 380-nm emission, because the FTP initially bound to the enzyme was only a small fraction (2–3%) of the total FTP in the cuvette.

ATP hydrolysed per mg protein min[-1] at 37 °C, and well over 99% of the activity was sensitive to ouabain. FTP binding to the enzyme was measured by mixing FTP and enzyme in the fluorimeter cell, in the presence of EDTA, and then measuring the reduction in fluorescence when the bound nucleotide was displaced by a large excess of ATP. Measurements of binding at several FTP concentrations gave the saturation curve shown in Fig. 1 and a calculated dissociation constant of 1.1 μM. A similar experiment with FDP gave a dissociation constant of 4.8 μM. ATP competed effectively with both FTP and FDP, and the calculated inhibition constant was 0.15 μM, which is similar to published values of 0.22 μM (ref. 7) and 0.12 μM (ref. 8) for the ATP dissociation constant.

Fluorescence enhancement could also be reversed by the addition of K[+], Tl[+], Rb[+], NH[4][+] and Cs[+], almost certainly because these ions cause dissociation of bound nucleotide[7]. Figure 2 shows the results of an experiment in which FDP was displaced from the enzyme by titration with KCl; virtually identical results were obtained with FTP. In the presence of 133 μM Na[+], 50% dissociation of nucleotide was obtained with a K[+] concentration of 140 μM.

The kinetics of FTP binding and FDP release during enzyme turnover were studied with the stopped-flow fluorimeter. Enzyme from one syringe was mixed with Mg[2+] and FTP from another, in the presence of Na[+] and Tris buffer. Figure 3 shows the results of two experiments with different time scales. Mixing was followed by (i) a very rapid rise in fluorescence, (ii) a rapid fall in fluorescence, with a rate constant of 14 s[-1] in Fig. 3a, (iii) a period of low fluorescence, and (iv) a final rise in fluorescence. The initial very rapid rise in fluorescence cannot be seen in Fig. 3 because of the time constant chosen. It almost certainly represents FTP binding, and it gives a calculated second-order rate constant of 2–8 × 10[7] M[-1] s[-1]; the rate was too fast to allow more precise determination. The rapid drop in fluorescence is most simply attributed to FDP release from
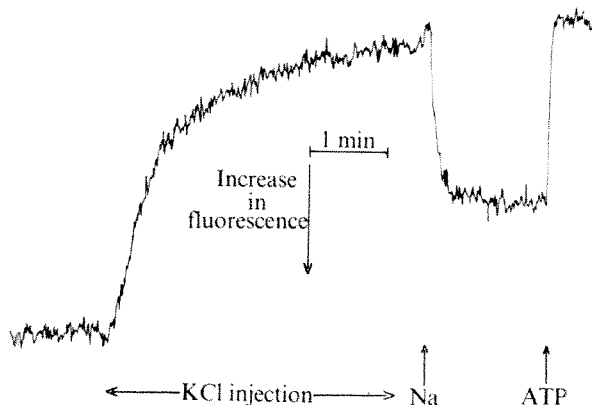
Fig. 2 Dissociation of FDP from $(Na^+ + K^+)$-ATPase by the addition of K^+ ions. The fluorimeter cell contained enzyme suspended at a concentration of 153 mg protein ml^{-1} in a solution containing FDP (triethylamine salt) 4 µM; Tris/Tris Cl (pH 7.7) 80 mM; EDTA (Tris salt) 1 mM; NaCl 133 µM. KCl was added from a motor-driven Agla syringe which delivered 5 µl of 150 mM KCl per minute. The syringe was stopped after 4 min, by which time the K^+ concentration was 1 mM. NaCl, to a final concentration of 17 mM, and ATP, to a final concentration of 27 µM, were added by hand. FTP fluorescence was excited at 310 nm and viewed at 380 nm as described in the legend to Fig. 1. An upward displacement of the trace represents a decrease in fluorescence and, because the rate of addition of KCl was very slow, the trace shows a succession of equilibrium states and gives no information about the rate of change of the enzyme (compare Fig. 4b).

the enzyme, though it could conceivably have been caused by a reduction in the fluorescence of bound FDP. The period of low fluorescence continued during hydrolysis of the FTP in the medium, and probably reflects the existence of a steady-state complex with a low affinity for nucleotide. The final rise in fluorescence may be attributed to binding of FDP from the medium as the concentration of FTP fell below its dissociation constant and the steady-state complex disappeared.

The presence of 1 mM K^+, added with the FTP, did not significantly alter the rate of fall of fluorescence (reflecting FDP release), though it did prolong the period of low fluorescence. The lack of effect of K^+ ions on the rate of fall of fluorescence is consistent with the Albers–Post sequence (see ref. 9), namely

$$E_1 + ATP \xrightleftharpoons[\quad]{Mg^{2+}, Na^+} E_1P.ADP \xrightleftharpoons[\quad]{\quad} E_1P + ADP$$

$$E_1P \xrightarrow{Mg^{2+}} E_2P \xrightarrow{K^+} P_i + E_2(K) \xrightarrow{ATP} E_1$$
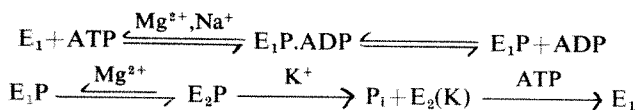
where E_1 is a form of the enzyme, with a high affinity for ATP, which exists in predominantly Na media, E_1P is a phosphorylated form that can transfer its phosphoryl group to ADP, E_2P is a phosphorylated form that is hydrolysed slowly spontaneously, but rapidly in the presence of K^+ ions, and $E_2(K)$ is a form of the enzyme (containing occluded K^+) that is produced by K^+-catalysed hydrolysis of E_2P and is converted to E_1 by a reaction which is accelerated by the binding of ATP at a low-affinity site. The low level of fluorescence during steady-state hydrolysis, both in the presence and absence of K^+ ions, also supports this scheme, by providing evidence that the release of ADP precedes the step that is rate limiting—presumably E_2P hydrolysis if K^+ ions are absent and $E_2(K)$ breakdown if they are present. The prolongation of the period of low fluorescence in the presence of K^+ is simply the result of the reduction in the overall hydrolysis rate, which occurs because the breakdown of $E_2(K)$ at these very low nucleotide concentrations is slower than the spontaneous hydrolysis of E_2P.

A series of experiments like that of Fig. 3b, but with different FTP concentrations, showed that the fall in fluorescence was progressively accelerated as the FTP concentration was increased

in the range 2–20 µM, beyond which measurements were impracticable.

We have also used changes in the affinity for FTP and FDP to investigate the rates of interconversion of the dephospho-enzyme between a form (E_2) with a low affinity for nucleotide, which exists when K^+ is the predominant cation, and a form (E_1) with a high affinity for nucleotide, which exists when Na^+ is the predominant cation[5,10-12]. Figure 4a shows the slow increase in fluorescence that occurred when enzyme preincubated with 0.8 mM K^+ in the absence of Na^+ was mixed with an equal volume of buffer containing 4 µM FTP and 100 mM Na^+. No magnesium was present and the buffer concentrations were adjusted to minimise changes in ionic strength. Since the binding of FTP to enzyme preincubated with Na^+ is very rapid (~ 200 s^{-1} at 4 µM FTP), the rate-limiting step in the experiment of Fig. 4a must have been the conversion of the K form to the Na form, or some process associated with that conversion.

In a series of similar experiments, increasing the FTP concentration in the range 4 to 24 µM increased the rate constant for the increase in fluorescence from ~ 0.25 to ~ 0.85 s^{-1}. Again, this points to the existence of a low affinity FTP binding site, and it suggests that binding of FTP at the low affinity site accelerates the change in conformation.

Change in conformation in the reverse direction (Na form to K form) was monitored in experiments like that of Fig. 4b, in which enzyme that had been loaded with nucleotide in the presence of 300 µM Na^+ was suddenly exposed to 1 mM K^+.
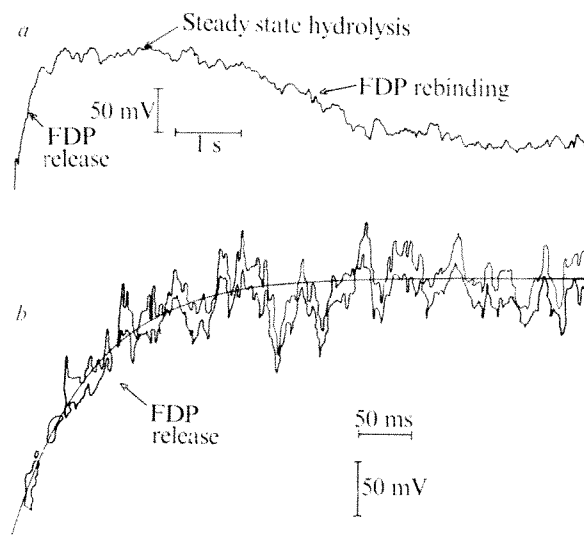
Fig. 3 Stopped-flow records of fluorescence changes accompanying FTP hydrolysis by $(Na^+ + K^+)$-ATPase. Experiments were done at 20–21 °C using a stopped-flow fluorimeter based on that described in ref. 14 but modified for split-beam operation. Fluorescence was excited at 310 nm, and the emitted light was detected after passage through Wratten 18A and Schott K.V.370 filters, which excluded most of the light below 360 nm. Note that an upward displacement of the trace represents a decrease in fluorescence. Syringe I contained: enzyme 300 µg protein ml^{-1}; NaCl 76 mM; Tris/Tris Cl (pH 7.7) 19 mM; histidine 1.1 mM; EDTA 40 µM. Syringe II contained: FTP (triethylamine salt) 8 µM in Fig. 3a and 14 µM in Fig. 3b; MgCl$_2$ 2.4 mM; NaCl 76 mM; Tris/Tris Cl (pH 7.7) 19 mM. The two syringes delivered equal volumes to the mixing chamber. a, Changes in fluorescence for a period of 10 s after mixing, recorded with a Datalab transient recorder. The time constant of the apparatus was set at 100 ms. b, Outline of an oscilloscope trace from a separate experiment showing changes in fluorescence for the first 0.5 s after mixing. The time constant was set at 5 ms. The continuous line represents the equation

$$y/y_\infty = 1 - \exp(-13.6t)$$

and is intended merely to show that, discounting noise, the fall in fluorescence can be described by a single exponential.
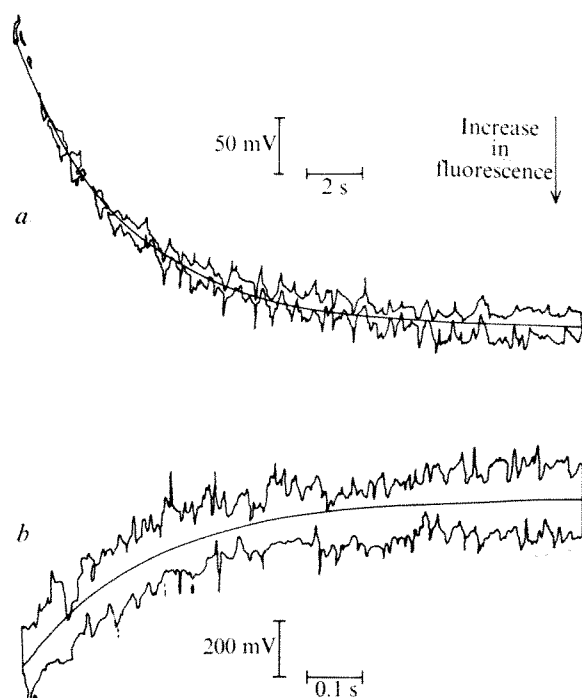
Fig. 4 Use of formycin nucleotides and the stopped-flow fluorimeter to follow slow conformational changes of the dephosphoenzyme induced by altering the concentrations of K^+ and Na^+ ions. *a*, Syringe I contained: enzyme 540 µg protein ml^{-1}; KCl 0.84 mM; Tris/Tris Cl (*p*H 7.7) 87 mM; histidine 4 mM; EDTA (Tris salt) 1 mM. Syringe II contained: FTP (triethylamine salt) 8 µM; NaCl 100 mM; Tris/Tris Cl (*p*H 7.7) 4 mM. The addition of Na^+ and FTP to the K form of the enzyme produced a slow increase in fluorescence, presumably caused by binding of FTP to the Na form of the enzyme as soon as it appeared. The time constant of the amplifier was 50 ms. The figure reproduces the outline of the oscilloscope trace, and the smooth curve represents the equation

$$y = y_0 \exp(-0.26t)$$

b, Syringe I contained: enzyme 380 µg protein ml^{-1}; FDP (triethylamine salt) 4 µM; NaCl 330 µM; Tris/Tris Cl (*p*H 7.7) 100 mM; EDTA (Tris salt) 1 mM; histidine 4 mM. Syringe II contained: FDP (triethylamine salt) 4 µM; KCl 2 mM; Tris/Tris Cl (*p*H 7.7) 100 mM; EDTA (Tris salt) 1 mM. The addition of K^+ to enzyme-FDP caused a slow fall in fluorescence, presumably the result of dissociation of bound FDP that occurred as soon as the enzyme was converted to the K form. The time constant of the apparatus was 1 ms. The figure reproduces the outline of the oscilloscope trace, and the smooth curve represents the equation

$$y/y_\infty = 1 - \exp(-4.4t)$$

In these experiments, FDP was used rather than FTP to avoid any possibility of phosphorylation, and its concentration was kept constant by including it in both syringes. Mixing was followed by a slow fall in fluorescence with a rate constant of about $4.4\ s^{-1}$. Since, in a control experiment (not shown), displacement of FDP from the Na form of the enzyme by the addition of excess ATP led to a rapid fall in fluorescence ($k = 110\ s^{-1}$), the slow fall in Fig. 4*b* presumably reflects a slow change from Na form to K form.

Although in the conditions of Fig. 4*b*, a K^+ concentration of 1 mM should have been sufficient to displace FDP almost completely (see Fig. 2), the rate of change of fluorescence was increased about fivefold when the K^+ concentration was increased from 1 mM to 5 mM. The amplitude of the change was unaffected. The increase in rate implies that binding of K^+ ions to a low affinity site accelerates the change in conformation, and suggests that the sequence of events is

$$E_1 + K^+ \rightleftharpoons E_1K \longrightarrow E_2K$$

where E_1, the form of the enzyme with a high affinity for nucleotide, has a low affinity for K^+. It is economical to suppose that E_2K in this equation is identical with '$E_2(K)$', the form of the enzyme with occluded K^+, in the scheme described above. The remarkably slow rates seen in the experiments like that of Fig.

4*a* are certainly compatible with the longevity of the 'occluded form' postulated by Post, Hegyvary and Kume[13] at low nucleotide concentrations. If this identification is correct, it follows that the occluded form can be reached not only by K^+-induced hydrolysis of phosphoenzyme, but also by conversion from a form of the enzyme that binds K^+ at low affinity (?intracellular) sites.

S. J. D. KARLISH

Biochemistry Department,
Weizmann Institute,
Rehovot, Israel

D. W. YATES

Department of Biochemistry,
University of Bristol,
Bristol, UK

I. M. GLYNN

Physiological Laboratory,
University of Cambridge,
Cambridge, UK

1 Ward, D. C., Cerami, A., Reich, E., Acs, G., and Altweger, L., *J. biol. Chem.*, 244, 3243–3250 (1969).
2 Ward, D. C., Reich, E., and Stryer, L., *J. biol. Chem.*, 244, 1228–1237 (1969).
3 Yount, R. G., *Adv. Enzymol.*, 43, 1–56 (1975).
4 Trentham, D. R., Eccleston, J. F., and Bagshaw, C. R., *Q. Rev. Biophys.*, 9 (in the press).
5 Jørgensen, P. L., *Biochim. biophys. Acta*, 401, 399–415 (1975).
6 Jørgensen, P. L., *Biochim. biophys. Acta*, 356, 36–52 (1974).
7 Hegyvary, C., and Post, R. L., *J. biol. Chem.*, 246, 5235–5240 (1971).
8 Nørby, J. G., and Jensen, J., *Biochim. biophys. Acta*, 233, 1004–1016 (1971).
9 Glynn, I. M., and Karlish, S. J. D., *A. Rev. Physiol.*, 37, 13–55 (1975).
10 Siegel, G. J., and Albers, R. W., *J. biol. Chem.*, 242, 4972–4979 (1967).
11 Post, R. L., Kume, S., and Rogers, F. N., in *Mechanisms in Bioenergetics* (edit. by Azzone, G. F., Ernster, L., Papa, S., Quagliariello, E., and Siliprandi, N.), (Academic, New York, 1973).
12 Skou, J. C., *Q. Rev. Biophys.*, 7, 401–434 (1975).
13 Post, R. L., Hegyvary, C., and Kume, S., *J. biol. Chem.*, 247, 6530–6540 (1972).
14 Gutfreund, H., *Prog. Biophys. molec. Biol.*, 29, 161–195 (1975).

General anaesthetics can selectively perturb lipid bilayer membranes

THE circumstantial evidence that anaesthetics act primarily by increasing the fluidity of membranes is quite strong. The gaseous, volatile, barbiturate, steroid and alcohol anaesthetics have all been shown to fluidise phosphatidylcholine–cholesterol lipid bilayers and the demonstration has also been made for some biological membranes[1–4]. Furthermore, a number of lipophilic substances, such as the higher alkanols, do not fluidise membranes and are not anaesthetics[3]. In some cases, a correlation between nerve-blocking potency and the action of anaesthetics in perturbing lipid bilayers has been observed[5]. Moreover, pressure counteracts the fluidising effects of anaesthetics just as it antagonises general anaesthesia *in vivo*[6–9]. The overall success of the fluidised lipid hypothesis tends to be its major drawback, for if anaesthetics fluidise membranes indiscriminantly then the hypothesis fails to provide a unique mechanism for their selective depression of neuronal function. We show here that lipid composition may modulate the ability of an anaesthetic to fluidise membranes more than has been generally supposed.

The effect of anaesthetics on membrane fluidity has been studied by intercalating fatty acids, or phosphatidylcholine (PC), labelled with a nitroxide reporter group into lipid bilayer membranes of various compositions. The anaesthetics studied showed wide variations in their ability to fluidise phospholipid bilayers. Results and experimental conditions are outlined in Table 1. All anaesthetics studied fluidised PC bilayers low in phosphatidic acid (PA) and high in cholesterol (Chol) (4%PA: 33%Chol) as evidenced by the consistent decrease in order parameter. At one extreme, halothane and urethane fluidised all the membranes studied, whereas at the other, pentobarbital and alphaxalone fluidised only those with 4%PA : 33%Chol.

Increasing cholesterol content conferred fluidising ability on all anaesthetics, but increasing the negatively charged PA tended to confer an ordering ability in some cases. A more detailed examination of the effects of lipid composition was made with pentobarbital and octanol (Table 2). The effect of pentobarbital switched from fluidising to ordering when, in the 4%PA : 33% Chol membranes, either the cholesterol content was lowered to 5–10% or PA was increased from 4 to 10%. Octanol always fluidised.

Our results in 4%PA : 33%Chol membranes are consistent with previous studies which have been carried out primarily in membranes of this cholesterol content, which is typical of nerve[2–4]. There are two isolated reports in the literature that halothane[11] and pentobarbital[12] order phospholipid membranes, as do local anaesthetics[12].

Spin-label[13] and deuterium magnetic resonance[14] studies suggest that the acyl groups of phospholipids are tilted near the head group region. General anaesthetics might reduce this tilt leading to a decrease in packing density as has been reported for tetracaine[13]. If so, the ordering effect should be weaker deeper in the bilayer and preliminary results with PC labelled with 8-doxylstearic acid at the β position are consistent with this explanation. In this context it is interesting to note that the three anaesthetics which increase anisotropy in the 4% PA bilayer all have structures which include rigid rings. A complete physical explanation of our results must await more detailed study, however.

The most important pharmacological aspect of this work concerns the specificity of action of anaesthetics. Although they are regarded conventionally as nonspecific drugs, many membrane processes are unaltered at anaesthetic doses. Thus the Na^+/K^+ ATPase of red blood cells[15] and synaptosomes[16] is unaffected at high, almost lytic, doses. This specificity might, *a priori*, reside in the primary perturbation of the lipid as well as in the secondary reaction of a given membrane protein to that lipid perturbation. Our work shows that the primary perturbation can no longer be thought of in terms of the lipid solubility of an anaesthetic alone; it is necessary to introduce the concept of fluidising efficacy, which we define as the rate of change of membrane fluidity with the concentration of anaesthetic in the membrane. Defined thus, fluidising efficacy

---

Table 1 The change in order parameter, ΔS, measured by 5-doxyl-stearic acid in lipid bilayer membranes exposed to anaesthetic agents

| Anaesthetic | Concentration | Bilayer composition, balance PC | | |
		4%PA	20%PA	4%PA:33% Chol
Halothane	11 mM	−0.03	−0.03	−0.06
Urethane	90 mM	−0.08	−0.03	−0.03
n-Octanol	46 mM	−0.02	−0.06	−0.08
Ketamine	25 mM	+0.03	0	−0.03
Alphaxalone	18 mM	+0.02	+0.03	−0.02
Pentobarbital	16 mM	+0.02	+0.02	−0.03

PC, Egg yolk phosphatidylcholine; PA, egg yolk phosphatidic acid; Chol, cholesterol; halothane is $CF_3CHClBr$; urethane is ethyl-carbamate; ketamine is 2-(methylamino)-2-(2-chlorophenyl) cyclo-hexanone and was a gift of Parke-Davis; alphaxalone is 3α-hydroxy-5α-pregnane-11,20-dione and was a gift of Glaxo. Lipids and spin labels in organic solvents were dried down together in pear-shaped flasks. Involatile anaesthetics were added in organic solvents before drying down; volatile agents were added later after the lipids had been dispersed in solution buffered at pH 7.0 by vortexing, and their concentration was checked by gas chromatography. Final lipid concentration was about 20–30 mg ml^{-1}, and the spin labels constituted about 1 mol % of the lipids. Anaesthetics were equilibrated with the bilayers up to 24 h before being sealed in 1-mm glass capillaries and equilibrated at 25 ± 0.5 °C in the cavity of either a Varian E-9 or E-109 electron spin resonance spectrometer operating at 9.5 GHz. Order parameters, ΔS, were calculated from spectra according to the method of Hubbell and McConnell[10]. A decrease in ΔS indicates a less anisotropic distribution of the label and a more fluid membrane. Changes in ΔS less than 0.01 were not considered significant. High doses of anaesthetics were used to obtain large changes in ΔS, which is a linear function of anaesthetic concentration within experimental error[2,3].

Table 2 The change in order parameter caused by anaesthetics in PC lipid bilayer membranes containing varying proportions of PA and Chol

| Anaesthetic | PA | Cholesterol | | | | |
		0	5%	10%	20%	33%
Pentobarbital	4%	+0.02	+0.02	0	−0.02	−0.03
16 mM	10%	—	+0.05	+0.04	+0.03	+0.05
	20%	+0.02	+0.09	+0.05	+0.02	+0.03
Octanol	4%	−0.02	—	—	—	−0.08
46 mM	10%	−0.05	−0.02	−0.03	−0.08	−0.04
	20%	−0.06	—	−0.03	−0.02	−0.03

---

need not be independent of anaesthetic concentration: thus Rosenberg reported unequivocal biphasic effects of halothane at physiological concentrations in palmitoyllauroyllecithin bilayers[11]. Most studies so far, however, have shown linear effects[2,3] or only weak nonlinear ones[4]. Three of the anaesthetics we examined could both order or fluidise bilayers, depending on their lipid composition, and thus indubitably exhibited both negative and positive fluidising efficacy; whereas for halothane, octanol and urethane, the fluidising efficacy was positive in all the membranes studied. For the latter anaesthetics we are unable to tell if the magnitude of the efficacy varies from membrane to membrane because the membrane solubilities are unknown. Our conclusions are thus based on the anaesthetics where a change in sign of the order parameter was observed.

To be consistent with the fluidised lipid hypothesis, anaesthetics should exhibit a positive fluidising efficacy at their site of action. In agreement with this, all six fluidised 4%PA: 33%Chol bilayers and pressure reversal of anaesthesia *in vivo* has been demonstrated for halothane[8,9], alphaxalone (as the clinical mixture althesin, which includes one third alphadolone acetate, a steroid with half the potency of alphaxalone)[9], pentobarbital[17], urethane[18] and ketamine (M. Wilson and K. W. M., unpublished). Long chain alcohols, which are not anaesthetics and do not fluidise membranes[3,19], might have lower efficacy than the short chain alcohols, or might simply have a lower membrane solubility. In the latter case, if the efficacy is about normal they might act additively with other anaesthetics as has been observed for some fluorinated hydrocarbons[20]. In all cases the magnitude of the fluidising efficacy can only be evaluated if the membrane partition coefficient is known, which is not often the case.

A corollary of the above argument is that only membrane regions where anaesthetics exhibit positive fluidising efficacy offer putative sites of action for anaesthetics. Thus, those anaesthetics, such as pentobarbital, which exhibit a positive fluidising efficacy in a restricted range of compositions might be useful for defining the nature of such sites. Our studies suggest they must have greater than 10% cholesterol and a charge density between 0.12 and 0.46 per phospholipid. This is consistent with the known compositions of neurones (30–50%Chol; 12–20% phospholipids bearing a single negative charge)[21], although our phospholipids are not typical of the variety found in biological membranes. Two factors could limit the usefulness of such an approach. First, the presence of protein in a lipid bilayer will probably exert an additional influence on the fluidising efficacy of anaesthetics–a single report that halothane orders rat brain synaptic membrane is cautionary[11]. Second, the lipids in biological membranes are heterogeneously distributed both across and in the plane of the membrane[22,23].

Large negative efficacies might also result in pharmacological effects[24]; indeed lipid ordering has been suggested as the cause of the hyperbaric convulsions exhibited by mammals at pressures in excess of 70 atm (ref. 25). We have shown that merely increasing the PA content from 4 to 10% in a 33% cholesterol bilayer changes the fluidising efficacy of pentobarbital from plus to minus. It remains to be seen how far the delicate balance

between anaesthetic and convulsant activities seen with many barbiturates can be explained in this manner.

Our work has been concerned exclusively with the primary anaesthetic-induced lipid perturbation. The sensitivity of membrane proteins to such perturbations might also vary.

KEITH W. MILLER
KAM-YEE Y. PANG

Departments of Pharmacology and Anesthesia,
Harvard Medical School and
Massachusetts General Hospital,
Boston, Massachusetts 02114

1 Metcalfe, J. C., Seeman, P., and Burgen, A. S. V., *Molec. Pharmac.*, 4, 87–95 (1968).
2 Trudell, J. R., Hubbell, W. L., and Cohen, E. N., *Biochim. biophys. Acta*, 291, 321–327 (1973).
3 Lawrence, D. K., and Gill, E. W., *Molec. Pharmac.*, 11, 280–286 (1975).
4 Boggs, J. M., Yoong, T., and Hsia, J. C., *Molec. Pharmac.*, 12, 127–135 (1976).
5 Johnson, S. M., Miller, K. W., and Bangham, A. D., *Biochim. biophys. Acta*, 307, 42–57 (1973).
6 Trudell, J. R., Hubbell, W. L., and Cohen, E. N., *Biochim. biophys. Acta*, 291, 328–334 (1973).
7 Johnson, S. M., and Miller, K. W., *Nature*, 228, 75–76 (1970).
8 Miller, K. W., Paton, W. D. M., Smith, R. A., and Smith, E. B., *Molec. Pharmac.*, 9, 131–143 (1973).
9 Halsey, M. J., and Wardley-Smith, B., *Nature*, 257, 811–813 (1975).
10 Hubbell, W. L., and McConnell, H. M., *J. Am. chem. Soc.*, 93, 314–326 (1971).
11 Rosenberg, P. H., Eibl, H., and Stier, A., *Molec. Pharmac.*, 11, 879–882 (1975).
12 Neal, M. J., Butler, K.W., Polnaszek, C. F., and Smith, I.C. P., *Molec. Pharmac.*, 12, 144–155 (1976).
13 McConnell, H. M., and McFarland, B. G., *Ann. N.Y. Acad. Sci.*, 195, 207–217 (1972).
14 Seelig, A., and Seelig, J., *Biochim. biophys. Acta*, 406, 1–5 (1975).
15 Hale, J., Keegan, R., Smith, E. B., and Snape, T. J., *Biochim. biophys. Acta*, 288, 107–113 (1972).
16 Levitt, J. D., *Anesthesiology*, 42, 267–274 (1975).
17 Lever, M. J., Miller, K. W., Paton, W. D. M., and Smith, E. B., *Nature*, 231, 368–371 (1971).
18 Johnson, F. H., and Flagler, E. A., *Science*, 112, 91–92 (1950).
19 Paterson, S. J., Butler, K. W., Huang, P., Labelle, J., and Smith, I. C. P., *Biochim. biophys. Acta*, 266, 597–602 (1972).
20 Miller, K. W., Paton, W. D. M., Smith, R. A., and Smith, E. B., *Anesthesiology*, 36, 339–351 (1972).
21 Whittaker, V. P., in *Handbook of Neurochemistry*, 2 (edit. by Lajtha, A.), 327–364 (Plenum, New York, 1969).
22 Gordesky, S. E., and Mannetti, G. V., *Biochem. biophys. Res. Commun.*, 50, 1027–1031 (1973).
23 Tanaka, K-I. and Ohnishi, S-I., *Biochim. biophys. Acta*, 426, 218–231 (1976).
24 Metcalfe, J. C., Hoult, J. R. S., and Colley, C. M., in *Molecular Mechanisms of General Anaesthesia* (edit. by Halsey, M. J., Miller, R. A., Sutton, J. A.), 145–163 (Churchill Livingstone, Edinburgh, 1974).
25 Miller, K. W., *Science*, 185, 867–869 (1974).

Restriction map of 5S RNA genes of *Drosophila melanogaster*

ALTHOUGH certain segments of the primary sequence and secondary structure of 5S RNA seem to be strikingly similar in prokaryotes and eukaryotes[1], the genetic organisation of the 5S RNA genes themselves exhibits remarkable variability. In *E. coli*, several lines of evidence indicate that the multiple copies of the 5S, 16S and 23S rRNA genes are organised into transcription units containing one gene each plus a sequence coding either for $tRNA_1^{ILE}$ or $tRNA_2^{GLU}$ such that they are cotranscribed in the order 16S–4S–23S–5S RNA (refs 2 and 3). In eukaryotes, the 5S RNA genes may be clustered at a single site as in humans[4,5] or widely distributed on many chromosomes as in the case of *Xenopus laevis*, where the 5S genes are found on the telomeres of most, if not all, of the 18 chromosomes[6]. Among the eukaryotes, there is no example in which the 5S genes are tightly linked to the 18S and 28S rRNA genes. Although there has been a large number of papers reporting the cytogenetic localisation of the 5S RNA genes in various organisms, only in the case of *Xenopus laevis* are there data concerning the physical organisation of these genes[7]. Even in this instance, however, the structure is only known for a few 5S gene repeat lengths, and the long range order that might be imposed on an entire 5S DNA cluster within the chromosome remains obscure. Whereas long range order in certain satellite sequences has been described[8,9], similar higher order organisation has not as yet been revealed for transcribed redundant genes.

In the fruit fly, *Drosophila melanogaster*, there are approxi-

mately 160 5S RNA genes (5S DNA) per haploid genome[10,11]. Both genetic[11] and *in situ*[12] hybridisation data indicate that these genes are clustered in the euchromatic portion of chromosome 2 at the cytogenetic locus 56EF. Utilising a series of genetic deficiencies, we have shown that it is possible to split 160 tandemly repeated 5S RNA genes into two functional clusters, each containing about 80 genes[11]. The results of experiments with restriction endonucleases reported here have allowed us to construct a high resolution restriction map depicting the physical organisation of these genes. The physical structure of the 5S DNA is similar to their genetic organisation in that the 160 genes are organised as two distinct clusters of 80 genes each.

On treatment of wild-type *D. melanogaster* DNA with *Eco*RI restriction endonuclease, the fragments are separated by electrophoresis on 1% Agarose slab gels and transferred to a cellulose nitrate filter according to the method of Southern[13]. The filter-bound fragments were hybridised with ^{125}I-labelled 5S RNA and finally visualised by autoradiography. As seen in Fig. 1, *D. melanogaster* 5S RNA hybridises to a single size class of 23 kb (kilobases or kilobase pairs) in length. Two independently isolated mutants (*min*) deficient for 50% of the wild-
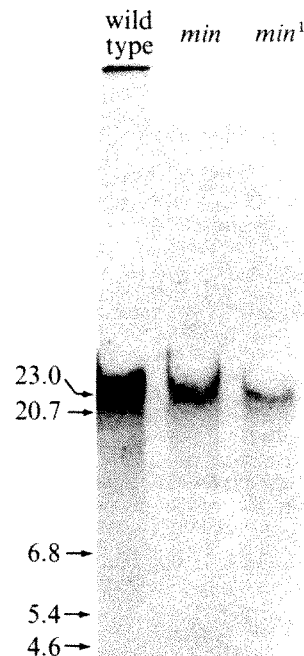
Fig. 1 Electrophoresis of *Eco*RI 5S DNA restriction fragments of wild-type and various *min* DNAs. *min* is a spontaneously derived mutant[11] and *min*[1] was obtained by triethylmalamine mutagenesis (Procunier and Tartof, unpublished). Each mutant contains half (0.003%) the number of wild-type 5S RNA genes. Approximately 10 µg of each DNA was digested with *Eco*RI in 0.05 ml of 50 mM Tris, *p*H 7.2, 50 mM NaCl, 10 mM $MgCl_2$ and 1 mM dithiothreitol at 37 °C for 2 h. A sufficient amount of enzyme was used to ensure complete digestion. The reaction was terminated by the addition of EDTA to 50 mM and NaCl to 0.20 M. The DNA was precipitated with 2 volumes of ethanol at −20 °C, collected by centrifugation and dissolved in E buffer (40 mM Tris, *p*H 7.2, 20 mM sodium acetate and 1 mM EDTA) containing 20% sucrose. Electrophoresis was carried out in 1% Agarose slab gels in E buffer for 4 h at 110 V. Following electrophoresis, the DNA fragments were transferred to a nitrocellulose filter according to the method of Southern[13] and then hybridised to ^{125}I-labelled 5S RNA (5 × 10⁷ c.p.m. µg⁻¹). The hybridisation reaction was carried out for 20 h at 60 °C by placing the filter in 10 ml of 2 × SSC (SSC is 0.15 M NaCl, 0.015 M sodium citrate, *p*H 7.0) with 0.20 µg of labelled 5S RNA and 400 times this amount of unlabelled 18S and 28S RNA. The filter was washed with 2 × SSC and treated with RNase (20 µg ml⁻¹) in 2 × SSC for 1 h at 37 °C. After another washing, the filter was dried at 60 °C and exposed to X-ray film (Ilford) for 13 d. *Eco*RI restriction fragments of phage λ DNA served as size standards (20.7, 6.8, 5.4 and 4.6 kb)[16]. From these, the size of the 5S DNA fragment was determined to be 23 kb.

Distance migrated (cm)
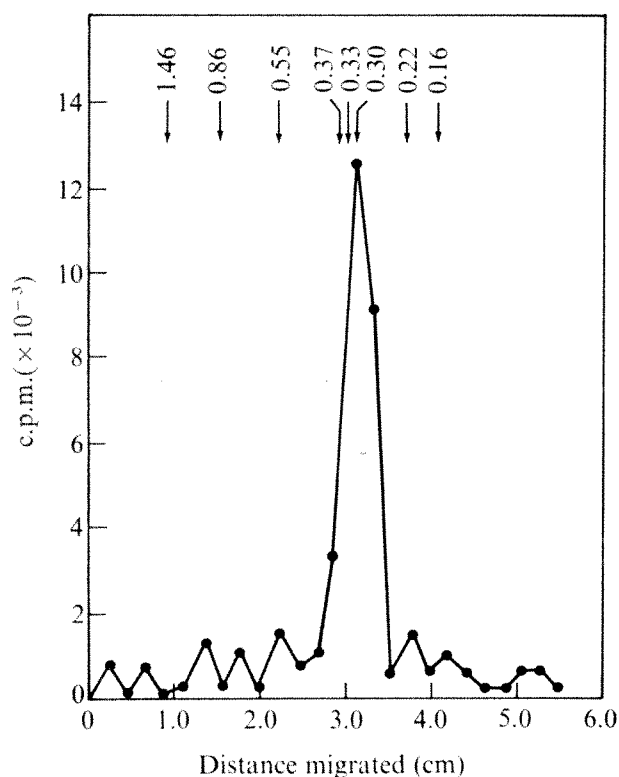
Fig. 2 Electrophoresis of *Hae*III 5S DNA restriction fragments. 400 μg of wild-type DNA were completely digested with purified *Hae*III at 37 °C in 6 mM Tris, *p*H 7.4, 6 mM NaCl, 6.6 mM MgCl₂ and 3 mM dithiothreitol. The enzyme was provided by Dr Jesse Summers of this Institute and purified according to Middleton *et al.*[17]. After digestion the reaction was terminated, DNA fragments precipitated and dissolved in E buffer according to the protocol of Fig. 1. Electrophoretic separation of the fragments was accomplished in 6% acrylamide —0.1% diacrylate gels (9.2 cm long, 0.45 cm diameter) in E buffer at 5 mA per gel for 1.5 h. The gel was frozen and sliced into 2-mm sections and each section dissolved in 50 μl of 1.5 N NaOH for 10 min at 70 °C. The mixture was neutralised with HCl and brought to a final concentration of 0.60 M NaCl, 0.20 M Tris and 0.02 M EDTA. The denatured DNA fragments were then covalently bound to phosphoimidazolol-cellulose filters according to Saxinger *et al.*[15]. The 5S ^{125}I-hybridisation reaction and subsequent RNase digestion was done as described previously (Fig. 1) and the filters dried and counted. Background values (identically treated filters but containing no DNA) were subtracted from the experimental values. *Hae*III restriction fragments of SV40 DNA provided the molecular weight standards in kilobases[18].

type number of 5S genes were similarly examined. The DNA from each of these also contains a 23-kb 5S DNA segment bounded by RI sites in spite of the fact that these mutants are deficient for half the wild-type number of these genes. This indicates there is more than one such 23-kb piece of 5S DNA per haploid genome.

To determine the number of *Eco*RI site bounded 5S DNA clusters, the precise size of the 5S repeating unit was obtained as follows. An *Hae*III cleavage point of 5S DNA can be predicted from the known *Drosophila* 5S RNA sequence[14]. The *Hae*III site

$$\begin{pmatrix} 5'\text{--GGCC--}3' \\ 3'\text{--CCGG--}5' \end{pmatrix}$$
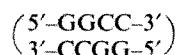
is between base pairs 116 and 119 of the 120 base pair 5S gene. The Southern technique which uses nitrocellulose filters however, fails to detect DNA fragments smaller than 500 base pairs[13]. Since it was quite possible that the 5S gene repeat size was smaller, we looked for a way round this difficulty. This was achieved by first separating the products of *Hae*III digestion by electrophoresis on a 6% polyacrylamide gel, hydrolysing the gel slices and denaturing the DNA fragments with alkali and then covalently binding the DNA fragments to a carbonyldi-

imidazole-derivatised phosphocellulose filter according to the procedure of Saxinger *et al.*[15]. The DNA-bound filters were then hybridised with ^{125}I-labelled 5S RNA. The results from this experiment are shown in Fig. 2, where a single *Hae*III fragment 290 base pairs long is found to hybridise to 5S RNA.

The results of the RI and *Hae*III restriction experiments suggest the following structure for the 5S DNA of *D. melanogaster*. 160 5S RNA genes with a repeat length of 290 base pairs would require 46.4 kb. The RI cuts embrace a cluster of 5S RNA genes 23 kb long. Therefore, we conclude that there are two RI site-bounded clusters, each containing 80 tandemly arranged 5S RNA genes. Each of these genes in turn consists of a 120 base pair structural gene and a 170 base-pair spacer. It is also known that 5S RNA is 57% GC, yet the genes band in a CsCl gradient in a position indicating that the average repeating unit has a base composition similar to the bulk DNA which is 40% GC (Fig. 3). So, we reason that the spacer is approximately 72% AT. A summary of the physical and genetic maps of *Drosophila* 5S DNA is given in Fig. 4.

It is possible that there could be a second *Hae*III site in the spacer region which would produce a second fragment that we would not detect in our experiments because the first *Hae*III site occurs three bases from the end of the 5S structural gene leaving an insufficient 5S DNA tag to probe for. The size of the *Hae*III fragments (290 base pairs) multiplied by the number of genes (160) however, fits almost exactly the dimension required to accommodate two RI site-bounded clusters (46.5 kb com-

Fig. 3 Buoyant density of 5S DNA. 100 μg of wild-type DNA were mixed with CsCl such that the final density was 1.700 g cm⁻³ in a volume of 7 ml. The solution was centrifuged for 50 h at 50,000 r.p.m. in a Beckman type 65 rotor at 20 °C. Fractions were collected and the absorption at 260 Å and buoyant density (ρ) measured. Each fraction was then denatured with alkali, neutralised and finally immobilised on a nitrocellulose filter[19]. The filters were then hybridised for 20 h at 60 °C in 3 ml of 2 × SSC containing 10⁶ c.p.m. of ^{125}I-labelled 5S RNA (5 × 10⁷ c.p.m. μg⁻¹) and 100 μg unlabelled 18S and 28S RNA. At the end of the incubation time, the filters were washed with 2 × SCC, treated with RNase (20 μg ml⁻¹) for 60 min at 37 °C, washed again with 2 × SSC and then dried and counted. ○, [125]I-5S RNA; ●, A_{260}; ———, ρ.

Fraction no.

Translocation breakpoints
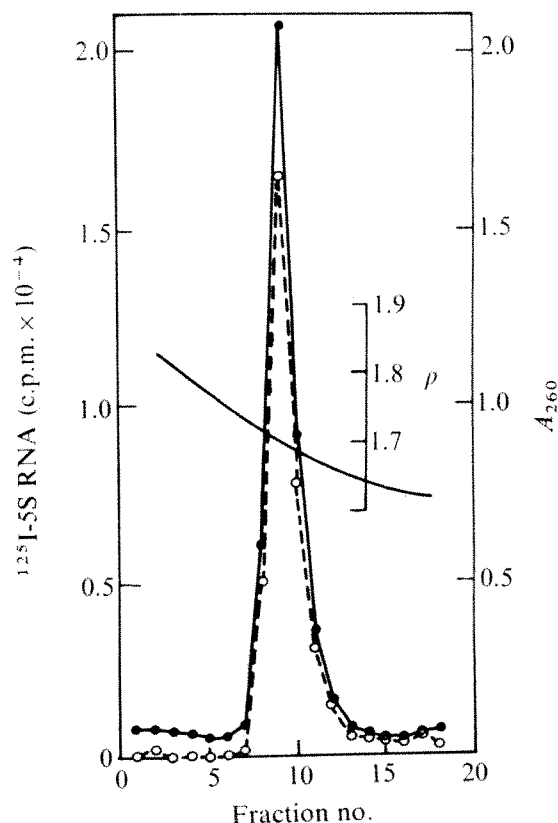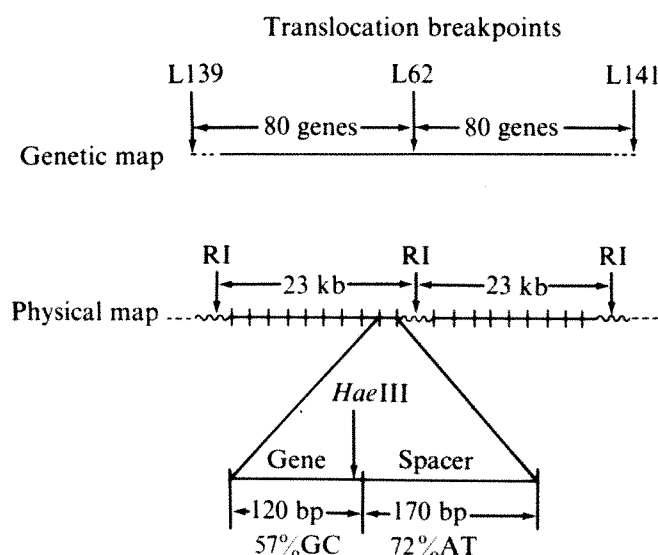
Fig. 4 A diagrammatic representation of the genetic and physical maps of the *Drosophila* 5S genes. From genetic evidence, it is known that the 5S DNA can be split into two functional clusters each of which contain approximately the same number of genes[11]. The designation and position of the translocation breakpoints which result in splitting the chromosome in the 5S DNA region are indicated. The physical map agrees with the genetic map in that there are two 5S DNA clusters of approximately equal size and bounded at the ends by *Eco*RI sites. The precise number of *Eco*RI sites and size of the DNA segment between the two 5S DNA clusters is unknown. The position of the *Hae*III site and the size and GC composition of structural gene and spacer are indicated.

pared with 46 kb), and this suggests to us that there is probably no other *Hae*III site in the spacer.

Two interesting aspects of the 5S DNA restrict n map deserve special comment. First, we do not know ei, er the quantity or the quality of the DNA segment inserted b 'ween the two 5S DNA clusters. The fact that one translocation breakpoint occurs between the two 5S DNA blocks (Fig. 4) suggests that the amount of DNA in this region must be appreciable. Second, the restriction map does not allow assignment of the relative order of the two tandemly arranged 5S DNA clusters with respect to each other. These clusters may be organised as direct serial repeats (a · b a · b) or as reverse repeats (a · b b · a). Unequal exchange between two direct serial repeats would be expected to produce genotypes containing variations in the number of RI site-bounded 5S DNA clusters. Indeed, one might expect the number of such clusters to increase enormously in the course of evolutionary time. Although our sample size is small, we have not observed such instability[11]. From such genetic considerations, we argue that a reverse repeat order of the two 5S DNA clusters would stabilise and preserve their structure in the chromosome. In addition, a palindrome arrangement would offer, on assuming the cruciform configuration, a means of maintaining homogeneity among both clusters through mismatch repair. We cannot but wonder if a similar arrangement might be imposed on other tandemly reiterated gene clusters.

JAMES D. PROCUNIER
KENNETH D. TARTOF*

The Institute for Cancer Research,
Fox Chase Cancer Center,
Philadelphia, Pennsylvania 19111

1 Fox, G. E., and Woese, C. R., *Nature*, 256, 505–507 (1975).
2 Pace, N. R., *Bact. Rev.*, 37, 562–603 (1973).
3 Lund, E., *et al.*, *Cell*, 7, 165–177 (1976).
4 Johnson, L. D., Henderson, A. S., and Atwood, K. C., *Cytogenet. Cell Genet.*, 13, 103–105 (1974).
5 Steffensen, D. H., Duffey, P., and Prensky, W., *Nature*, 252, 741–743 (1974).
6 Pardue, M. L., Brown, D. D., and Birnstiel, M. L., *Chromosoma*, 42, 191–203 (1973).
7 Carroll, D., and Brown, D. D., *Cell*, 7, 467–475 (1976).
8 Botchan, M., *Nature*, 251, 288–292 (1974).
9 Southern, E. M., *J. molec. Biol.*, 94, 51–69 (1975).
10 Tartof, K. D., and Perry, R. P., *J. molec. Biol.*, 51, 171–183 (1970).
11 Procunier, J. D., and Tartof, K. D., *Genetics*, 81, 515–523 (1975).
12 Wimber, D. E., and Steffensen, D. M., *Science*, 170, 639–641 (1970).
13 Southern, E. M., *J. molec. Biol.*, 98, 503–517 (1975).
14 Benhamou, J., and Jordan, B. R., *FEBS Lett.*, 62, 146–149 (1976).
15 Saxinger, W. C., Ponnamperuma, C., and Gillespie, D., *Proc. natn. Acad. Sci. U.S.A.*, 69, 2975–2978 (1972).
16 Allet, B., Jeppesen, P. G. N., Katagiri, K. J., and Delius, H., *Nature*, 241, 120–123 (1973).
17 Middleton, J. H., Edgell, M. H., and Hutchinson, C. A., III, *J. Virol.*, 10, 42–50 (1972).
18 Lebowitz, P., Siegel, W., and Sklar, J., *J. molec. Biol.*, 88, 105–123 (1974).
19 Birnstiel, M., Speirs, J., Purdom, I., and Jones, K., *Nature*, 219, 454–463 (1968).

Two proteins function in the regulation of photosynthetic CO_2 assimilation in chloroplasts

THERE is increasing evidence that products of the light reactions of photosynthesis govern the activity of enzymes involved in CO_2 assimilation by chloroplasts[1–15]. Of these products, reductants formed photochemically seem to be of particular importance. Such reductants include the reduced form of ferredoxin[2,6,16], a strongly electronegative chloroplast iron–sulphur protein ($E'_0 = -0.42$ V) that activates the two key chloroplast enzymes fructose 1,6-*bis*-phosphatase and sedoheptulose 1,7-*bis*-phosphatase. Activation of both of these enzymes requires in addition to reduced ferredoxin a 'protein factor' that is indigenous to chloroplasts. In efforts to elucidate the nature of the ferredoxin-linked enzyme activation, we have separated the protein factor into two components[16]: (1) a partly purified protein, provisionally named "assimilation regulatory protein a" (ARP_a) and (2) a highly purified chromophore-free protein called "assimilation regulatory protein b" (ARP_b). Only the latter was required for activation when reduced ferredoxin was replaced by the non-physiological sulphydryl reagent dithiothreitol[6,19].

The protein factor was resolved into the ARP_a and ARP_b components by following the reported purification procedure (fractionation by acid pH, acetone, ammonium sulphate, and Sephadex G-100 gel filtration)[6] that was modified to provide improved resolution in the gel filtration step.

Table 1 Requirement for ARP_a and ARP_b for activation of chloroplast fructose 1,6-*bis*-phosphatase by reduced ferredoxin

	nmol P_i released per min
Control	0
ARP_a	14
ARP_b	0
ARP_a plus ARP_b	73

The reaction was carried out in Warburg vessels containing (in the sidearm) 6 μmol of sodium fructose 1,6-*bis*-phosphate, and (in the main compartment) 135 μg of fructose 1,6-*bis*-phosphatase; spinach chloroplast fragments, P_1S_1, equivalent to 0.1 mg chlorophyll[6]; 0.12 mg of spinach ferredoxin, and the following (μmol): Tris-HCl buffer, pH 8.0, 100; neutralised reduced glutathione, 5; sodium ascorbate, 10; 2,6-dichlorophenolindophenol, 0.1; $MgCl_2$, 1.85 μg of ARP_a and 108 μg of ARP_b, were added as indicated. Vessels were equilibrated with nitrogen for 6 min and were then incubated for 5 min in the light. The reaction was started by adding fructose 1,6-*bis*-phosphate from the sidearm and was continued for 30 min under illumination. Light intensity, was 20,000 lx and temperature was 20 °C. The reaction was stopped by adding 0.5 ml of 10% trichloroacetic acid and (after removing the precipitate by centrifugation) P_i was estimated by a modified Fiske–SubbaRow procedure[6].

Although ARP_a and ARP_b were not completely separated, even after improvement in the gel filtration step, fractions of the two components were sufficiently resolved to establish their separate identities. ARP_b was then further purified by DEAE–cellulose chromatography and a second Sephadex G-100 gel filtration step. Details of the additional purification procedure will be published elsewhere.

Ultracentrifugation of purified ARP_b (single absorption maximum at 280 nm) showed one component with an approximate molecular weight of 20,000 but polyacrylamide gel electrophoresis yielded two components, both of which showed ARP_b activity. It seems therefore that ARP_b can exist in more than one form.

Table 1 demonstrates that both ARP_a and ARP_b were required for activation of fructose 1,6-*bis*-phosphatase by reduced ferredoxin. By contrast (Table 2), only ARP_b was required for activation of the enzyme when reduced ferredoxin was replaced by dithiothreitol. Similarly, the dithiothreitol-induced activation of the sedoheptulose 1,7-*bis*-phosphatase activity that occurs when the chloroplast fructose 1,6-*bis*-phosphatase is converted from its dimer to its monomer form when the pH is increased from 5.5 to 8.5 (ref. 17) required only ARP_b.

In addition to its activation of the fructose 1,6-*bis*-phosphatase and sedoheptulose 1,7-*bis*-phosphatase enzymes of chloroplasts, ARP_b was found to activate the regulatory form of NADP–glyceraldehyde phosphate dehydrogenase[14,18] in the presence of dithiothreitol (Table 3). Neither the NAD-linked activity of the regulatory form of the enzyme nor the NAD(P)-linked activity of its non-regulatory counterparts was affected by ARP_b. In the presence of dithiothreitol, ARP_b also activated one of the two forms of phosphoribulokinase that we have found in chloroplasts (Table 3). In neither case could other sulphydryl reagents replace dithiothreitol.

In summary, the evidence indicates that ARP_a and ARP_b, two newly discovered chloroplast proteins, link the strong reducing power of photochemically reduced ferredoxin to the regulation of certain enzymes of photosynthetic CO_2 assimilation. We believe that the ferredoxin-linked system constitutes a specific mechanism of enzyme regulation that operates in addition to the other light-actuated systems of chloroplasts. Of the latter group, those mechanisms involving enzyme effectors (such as reduced NADP) would govern

Table 3 Effect of preincubation with ARP_b and dithiothreitol on chloroplast NADP–glyceraldehyde phosphate dehydrogenase and phosphoribulokinase

	$\Delta A_{340\ nm}$ per min Preincubation with	
	dithiothreitol	ARP_b plus dithiothreitol
NADP–glyceraldehyde phosphate dehydrogenase	0.30	1.14
Phosphoribulokinase	0.04	0.37

To assay NADP–glyceraldehyde phosphate dehydrogenase, 90 µg of the regulatory form of the enzyme[14] was preincubated for 3 min in a solution containing (in a volume of 0.1 ml) 10 µmol of Tricine–NaOH buffer, pH 8.4; 1 µmol of dithiothreitol, and, as indicated, 9 µg of ARP_b. After preincubation, the mixture was injected into the assay mixture that contained, in a volume of 0.9 ml, 3 µg of 3-phosphoglycerate phosphokinase and the following (µmol): Tricine–NaOH buffer, pH 8.4, 40; dithiothreitol, 1.5, $MgCl_2$, 10; ATP, 5; 3-phosphoglycerate, 5; NADPH, 0.12. Enzyme activity was measured by following the change in absorbance at 340 nm. The temperature was 22 °C. To assay phosphoribulokinase[24] (the form present in the purified fraction containing the regulatory form of NADP–glyceraldehyde phosphate dehydrogenase[14]), 90 µg of the enzyme was preincubated for 3 min in a solution containing (in a volume of 0.1 ml) 10 µmol of Tris-HCl buffer, pH 7.5; 1 µmol of dithiothreitol, and, as indicated, 9 µg of ARP_b. After preincubation, the mixture was injected into the assay solution that contained, in a volume of 0.9 ml, 10 µg of phosphoribuloisomerase; 60 µg of lactate dehydrogenase; 30 µg of pyruvate kinase, and the following (µmol): Tris-HCl buffer, pH 7.5, 100; reduced glutathione, 10; $MgCl_2$, 10; ribose-5-phosphate, 2; phosphoenolpyruvate, 0.5; ATP, 0.5; NADH, 0.12. Enzyme activity was measured by following the change in absorbance at 340 nm. The temperature was 22 °C.

the activity of selected regulatory enzymes, whereas other mechanisms (for example, those involving light-induced increases in the pH (refs 19, 20) or Mg^{2+} concentration[21,23] of the stroma or soluble phase of chloroplasts) would provide the environment necessary for the optimal activity of enzymes of the reductive pentose phosphate cycle.

PETER SCHÜRMANN
RICARDO A. WOLOSIUK
VICKI D. BREAZEALE
BOB B. BUCHANAN

*Department of Cell Physiology,
University of California,
Berkeley, California 94720*

Table 2 Activation of the fructose 1,6-*bis*-phosphatase and sedoheptulose 1,7-*bis*-phosphatase activities of chloroplast fructose 1,6-*bis*-phosphatase by ARP_b and dithiothreitol

Treatment	nmol of P_i released per min Fructose 1,6-*bis*-phosphatase	Sedoheptulose 1,7-*bis*-phosphatase
Complete	168	131
Minus $MgCl_2$	7	1
Minus ARP_b	14	58
Minus fructose 1,6-*bis*-phosphatase	6	4
Minus dithiothreitol	11	4
Minus dithiothreitol, plus reduced glutathione	10	4
Minus dithiothreitol, plus β-mercaptoethanol	11	5

To assay fructose 1,6-*bis*-phosphatase activity, the complete system contained 8 µg of the dimer form of fructose 1,6-*bis*-phosphatase[17], 9 µg of ARP_b, and the following (µmol): Tris-HCl buffer, pH 8.0, 100; $MgCl_2$, 1; dithiothreitol, 5; and fructose 1,6-*bis*-phosphate, 6. The final volume was 1.0 ml. To assay sedoheptulose 1,7-*bis*-phosphatase, the complete system contained 16 µg of the monomer form of fructose 1,6-*bis*-phosphatase (obtained from the enzyme dimer by increasing the pH from 5.5 to 8.5 (ref. 17); 9 µg of ARP_b, and the following (µmol): Tris-HCl buffer, pH 8.0, 100; $MgCl_2$, 10; dithiothreitol, 5; sedoheptulose 1,7-*bis*-phosphate, 1.5. The final volume was 1.0 ml. In each case, the reaction (at 25 °C) was started by the addition of enzyme and stopped by the addition of 4 ml of the mixture used for P_i analysis (see Table 1).

1 Petersen, T. A., Kirk, M., and Bassham, J. A., *Physiologia Pl.*, **19**, 219–231 (1966).
2 Buchanan, B. B., Kalberer, P. P., and Arnon, D. I., *Biochem. biophys. Res. Commun.*, **29**, 74–79 (1967).
3 Bassham, J. A., Kirk, M., and Jensen, R. G., *Biochim. biophys. Acta*, **153**, 211–218 (1968).
4 Müller, B., Ziegler, I., and Ziegler, H., *Eur. J. Biochem.*, **9**, 101–106 (1969).
5 Latzko, E., Garnier, R. V., and Gibbs, M., *Biochem. biophys. Res. Commun.*, **39**, 1140–1144 (1970).
6 Buchanan, B. B., Schürmann, P., and Kalberer, P. P., *J. biol. Chem.*, **246**, 5952–5959 (1971).
7 Buchanan, B. B., and Schürmann, P., *J. biol. Chem.*, **248**, 4956–4964 (1973).
8 Pupillo, P., and Piccari, G. G., *Archs. Biochem. Biophys.*, **154**, 324–331 (1973).
9 Chu, D. K., and Bassham, J. A., *Pl. Physiol.*, **54**, 556–559 (1974).
10 Schürmann, P., and Buchanan, B. B., *Biochim. biophys. Acta*, **376**, 189–192 (1975).
11 Champigny, M. L., and Bismuth, E., *Physiologia Pl.*, **36**, 95–100 (1976).
12 Ryan, F. J., and Tolbert, N. E., *J. biol. Chem.*, **250**, 4234–4238 (1975).
13 Anderson, L. E., and Avron, M., *Pl. Physiol.*, **57**, 209–213 (1976).
14 Wolosiuk, R. A., and Buchanan, B. B., *J. biol. Chem.* (in the press).
15 Marsho, T. V., and Kung, S. D., *Archs. Biochem. Biophys.*, **173**, 341–346 (1976).
16 Wolosiuk, R. A., Schürmann, P., Breazeale, V. D., and Buchanan, B. B., *Fedn Proc.*, **35**, 1732 (1976).
17 Buchanan, B. B., Schürmann, P., and Wolosiuk, R. A., *Biochem. biophys. Res. Commun.*, **69**, 970–978 (1976).
18 Pupillo, P., and Giuliani Piccari, G., *Eur. J. Biochem.*, **51**, 475–482 (1975).
19 Heldt, H. W., Werdan, K., Milovancev, M., and Geller, G., *Biochim. biophys. Acta*, **314**, 224–241 (1973).
20 Werdan, K., Heldt, K. W., and Milovancev, M., *Biochim. biophys. Acta*, **396**, 276–292 (1975).
21 Lin, D. C., and Nobel, P. S., *Archs. Biochem. Biophys.*, **145**, 622–632 (1971).
22 Lilley, R. McC., Holborow, K., and Walker, D. A., *New Phytol.*, **73**, 657–662 (1974).
23 Barber, J., *Trends biochem. Sci.*, **1**, 33–36 (1976).
24 Moll, B., and Levine, R. R., *Pl. Physiol.*, **46**, 576–580 (1970).

matters arising

Biotic extinctions by solar flares

REID *et al.*[1] have suggested a mechanism by which solar protons might catastrophically deplete atmospheric ozone during a reversal of the Earth's geomagnetic field, when its shielding effect is weakened. Organisms would thereby be exposed to a harsher ultraviolet environment, producing extinctions, such as those Hays[2] observed, closely correlated with geomagnetic reversals in deep-sea sediment cores. They further suggest that mass extinctions, such as those which took place at the close of the Cretaceous, may have thus occurred. Reid *et al.* assume that during a reversal the geomagnetic field effectively disappears for $\sim 1,000$ yr. They also assume that solar flares sufficiently intense to cause extinctions occur at intervals of $\sim 1,000$ yr or more. We propose to examine the validity of these assumptions by comparing them with geomagnetic reversals identified by Tarling and Mitchell[3] for the past 75 Myr, and small scale (radiolarian) and large scale (dinosaurian) extinctions.

Hays[2] found four radiolarian cases of extinction (involving 6 species) during the past 2.5 Myr, in which time 10 geomagnetic reversals occurred[3]. According to the model proposed by Reid *et al.*, the expected number of extinction events (E) is related to the probability of occurrence of a strong solar flare during any year (P_s) and the number of successive years during which the magnetic field is effectively absent (T) as

$$E = RTP_s \qquad (1)$$

where R is the number of reversals during the period considered. Solving for P_s, and substituting the values cited above:

$$P_s = E/RT = 4 \times 10^{-4} \text{ yr}^{-1} \qquad (2)$$

The value for P_s is lower than assumed by Reid *et al.* Given their estimate of $P_s = 10^{-3} \text{ yr}^{-1}$, to satisfy equation (1), the magnetic field would have to disappear for 400 yr during a reversal, an interval shorter than the generally accepted value[4]. A reduction in P_s would not, in our view, impair the utility of the model proposed by Reid *et al.* in producing periods of small scale extinction.

It is, however, evident that the terminal Cretaceous (dinosaurian) extinctions affected a much broader range of organisms[1,5] than the radiolarian ones described by Hays. The extinction of the dinosaurs occurred within a short but undefined interval ~ 18 Myr after the end of a lengthy period of normal polarity[6,7]. During these 18 Myr, eleven reversals took place[6], but the diversity of terrestrial and marine reptiles remained constant up to the end of the Cretaceous[8].

A solar flare powerful enough to produce such extinctions would, therefore, have been much more powerful than any directly considered by Reid *et al.* Because no extinction as severe as those in which the dinosaurs were eliminated have subsequently taken place (or $E = 1$), and at least 197 reversals have since occurred ($R = 197$; set $T = 10^3$ yr), then the frequency of a hypothetical giant flare according to equation (2) is

$$P_s = (1.97 \times 10^5 \text{ yr})^{-1}$$

or ~ 1 per 200,000 yr. Solar proton events have been studied only during the past 15–20 yr (ref. 1), and there is little hard evidence for the existence of giant flares, although instabilities in solar activity are now receiving more attention than formerly[9]. We therefore concur that the supernova model, as proposed by Terry and Tucker[10] and Ruderman[11] remains worthy of consideration. Although the coincidence of a reversal and the terminal Cretaceous extinctions would invalidate neither model, Keating[12] notes that this coincidence has not been demonstrated.

P. BÉLAND
D. A. RUSSELL

National Museum of Natural Sciences, Ottawa, Canada

1 Reid, G. C., Isaksen, I. S. A., Holzer, T. E., and Crutzen, P. J., *Nature*, 259, 177–179 (1976).
2 Hays, J. D., *Bull. geol. Soc. Am.*, 82, 2433–2447 (1971).
3 Tarling, D. H., and Mitchell, J. G., *Geology*, 4, 133–136 (1976).
4 Dunn, J. R., Fuller, M., Ito, H., and Schmidt, V. A., *Science*, 172, 840–845 (1971).
5 Russell, D. A., and Tucker, W. H., *Nature*, 229, 553–554 (1971).
6 Lowrie, W., Alvarez, W., and Premoli Silva, I., *EOS, Trans. Am. geophys. Un.*, 57, 238 (1976).
7 Obradovich, J. D., and Cobban, W. A., *Geol. Ass. Can., Special Paper*, 13, 31–54 (1975).
8 Russell, D. A., *Geol. Ass. Can., Special Paper*, 13, 119–136 (1975).
9 Hammond, A. L., *Science*, 191, 1159–1160 (1976).
10 Terry, K. D., and Tucker, W. H., *Science*, 159, 421–423 (1968).
11 Ruderman, M. A., *Science*, 184, 1079–1081 (1974).
12 Keating, B., Helsley, C. E., and Pessagno, E. A., Jr, *Geology*, 3, 73–76 (1975).

CRUTZEN AND REID REPLY—We certainly agree with the interesting remarks of Béland and Russell[1]. Our model was primarily intended to provide a tentative explanation for the apparent mysterious association between geomagnetic polarity reversals and small scale extinctions, as documented by Hays[2]. The relationship between polarity reversals and mass extinctions, such as that at the close of the Cretaceous, is not well established, and our mechanism is only a possible candidate.

In terms of ionising radiation, the solar flares of August 1972 dissipated $\sim 6 \times 10^5 \text{ erg cm}^{-2}$ in the polar stratosphere. A similar flux of energy over the entire projected area of the Earth would be caused by the γ-ray 'pulse' ($\sim 10^{49}$ erg (ref. 3)) from a supernova 1,000 light yr away, while a supernova 30 light yr away would create a flux 1,000 times larger. It is not likely that the Sun could produce such a colossal flare, but several supernovae may have occurred within a distance of 30 light yr during the lifetime of the Solar System[4].

Although the shielding of ultraviolet light by NO_2 must be considered, any ozone depletion caused by such supernovae would seriously affect the biosphere[5]. However, other consequences of the vast amounts of nitrogen fixed in the atmosphere are worth considering. Assuming maximum efficiency of NO production, the supernova would fix $\sim 1,500$ Mt of nitrogen, which is almost 10 times the presently estimated annual global nitrogen fixation rate[6]. This could perturb ecological systems. Furthermore, the column density of NO_x could initially reach unhealthy concentrations of 10^{19}–10^{20} molecules per cm^2 and although detailed calculations are necessary to confirm the idea, it may seem that there could be substantial production of NO_2 and even O_3 by the reaction $2NO + O_2 \rightarrow 2NO_2$. Given absorption cross sections of $NO_2 > 10^{-19} \text{ cm}^2$ at wave lengths $\lesssim 650$ nm, severe perturbations in photosynthesis rates and in the radiation balance of the Earth may result.

Aeronomy Laboratory,
NOAA, Boulder,
Colorado 80302

1 Béland, P., and Russell, D. A., *Nature*, 263, 259 (1976).
2 Hays, J. D., *Bull. geol. Soc. Am.*, 83, 2433 (1971).
3 Tucker, W. H., and Terry, K. D., *Science*, 160, 1138 (1968).
4 Shklovsky, I. S., *Supernovae* (John Wiley and Sons, London, 1968).
5 Ruderman, M. A., *Science*, 184, 1079 (1974).
6 Hardy, R. W. F., and Havelka, U. D., *Science*, 188, 633 (1975).

Glaciations and dense interstellar clouds

As a possible cause of an ice age on Earth, I have suggested the passage of the Solar System through a region of compression of interstellar matter (ISM) bordering a spiral arm of the Galaxy[1]. Dennison and Mansfield criticise the model because it leads them to expect to find a dense cloud of ISM still very close to us; no such cloud is seen[2]. Their criticism, however, ignores the structure of the Galaxy that provides the basis for my suggestion, and the justification for reviving the idea of a possible relation between ISM and ice ages.

The compression region is a shock region, that is, a 'traffic jam', in the ISM. According to the model, the most recent glaciation was associated with the Sun having been immersed in a cloud of ISM while it traversed such a region. On my interpretation, the Sun emerged from the region after that. Dennison and Mansfield adopt ~ 20 km s^{-1} for the relative speed of the Sun and cloud, and $\sim 10^4$ yr ago for the time of emergence.

Dennison and Mansfield take no account of the fact that the Sun and cloud were travelling through the region with mean speed ~ 250 km s^{-1} in their orbital motion round the Galaxy. So the place where the Sun emerged from the cloud is more than 10 times further off than in the reckoning of Dennison and Mansfield.

The material of the cloud concerned did indeed emerge from the compression region about the same time, but there are three essential points to recognise: (1) The region forms an oblique shock and, viewed as such, it produces a discontinuity in direction of flow of the ISM passing through it, but not of the stars. After emerging from the region, the cloud material and the Sun have relative motion very different from what they had in it. The traffic-jam is like congestion on an escalator—after people get off an escalator they disperse with relative speeds comparable to the speed of the escalator, not to the relative speed they had while on the escalator. (2) When the material emerges it is no longer compressed, since by definition the compression region is the region where it is compressed. Compression regions themselves are common places of observation, but how the compressed material evaporates out of them is not understood in detail. (3) There is no reason to suppose that clouds retain their individuality as the ISM evaporates from a compression region.

The answer to Dennison and Mansfield is that the dense cloud they expect is not where they infer, it is not dense, and it is not a cloud.

For reasons I have stated briefly elsewhere[1,3], the end of the most recent glaciation could have been the very end of a period of glacial activity that did not even start until after the Sun had emerged from the cloud, and its luminosity had fallen back to normal after an interval of enhancement. Thus the time, $\sim 10^4$ yr, is almost certainly an underestimate. In their closing paragraph Dennison and Mansfield go some way towards admitting this further general consideration and its implications. I should take issue with them over several other particular matters were this necessary. It seems, to me, however, that their objections vanish in the light of the foregoing general considerations, and that their discussion actually helps to show that there need be no embarrassing side effects in the model.

I was aware of the problem of a nearby cloud raised by Hoyle and Lyttleton in a paper[4] I quoted in ref. 1, and recalled by Dennison and Mansfield, but ideas on the Galaxy and the kinematics of ISM have changed since 1939.

W. H. McCrea

Astronomy Centre,
University of Sussex,
Falmer, Sussex BN1 9QH, UK

[1] McCrea, W. H., *Nature*, 255, 607–609 (1975).
[2] Dennison. B., and Mansfield, V. N., *Nature*, 261, 32–34 (1976).
[3] McCrea. W. H., *Observatory*, 95, 239–255 (1975).
[4] Hoyle, F., and Lyttleton, R. A., *Proc. Camb. Phil. Soc.*, 35, 405–415 (1939).

DENNISON AND MANSFIELD REPLY—On encountering a spiral density wave, interstellar matter (ISM) suffers rapid and substantial changes in its flow field and density. Behind the shock, the gas gradually recovers over some millions of years[1]. The actual shock region is extremely narrow, and thus the transit time of the stars and gas through it is negligible in comparison to a million years, the approximate duration of the last general ice age. Behind the shock, the ISM, or any cloud, could not undergo significant changes in velocity, density, or identity caused by spiral density wave in the $\sim 10^4$ yr since the last glaciation, or even in the $\sim 7 \times 10^4$ yr since the beginning of the last glaciation. To accomplish the rapid changes now proposed by McCrea[2], an unjustified assumption of a second abrupt change in the flow field of the ISM behind the initial shock would have to be made. The compression region fades gradually—it does not end abruptly like people leaving an escalator.

Before entering the compression region, the stars and ISM have a small velocity differential of ~ 10 km s^{-1}, because they follow similar orbits. In entering the compression region, the ISM is severely perturbed, while the stars are largely unaffected. This produces velocity differentials between the stars and gas that are an order of magnitude larger than the 5–25 km s^{-1} that McCrea uses. His low velocity differentials are appropriate for regions outside the compression zone. Since the 'traffic jam' applies to the ISM only, it is in the compression zone that there are large velocity differentials, and because the accretion rate in the model depends linearly on the cloud density and inversely as the cube of the velocity differential, the same fractional increase in solar luminosity would now require the frequent encountering of clouds with embarrassingly high densities, from $\sim 10^8$ to $\sim 10^{10}$ hydrogen molecules cm^{-3}.

Even if we accept the unorthodox picture that McCrea now uses, an excessive amount of energy is needed to dissipate these dense clouds on time scales of $\sim 10^4$ yr. A typical cloud required by the original model has $\sim 10^6$ hydrogen molecules cm^{-3} and a diameter of ~ 1 pc. Such dense clouds cool rapidly and cannot store compressional energy. It is easy to show that over a time of $\sim 10^4$ yr a power of $\sim 5 \times 10^4 L_\odot$ is required to overcome the gravitational binding of the cloud. McCrea's present model would require the dissipation of many such clouds throughout the disk of the Galaxy. This requires that on the average a power $\gtrsim$ total luminosity of the Galaxy must be constantly expended in dissipating dense clouds.

Center for Radiophysics and
Space Research,
Cornell University
Department of Physics and Astronomy,
Colgate University

[1] Woodward, P. R., *Astrophys. J.*, 195, 61–73 (1975).
[2] McCrea, W. H., *Nature*, 263, 260 (1976).

reviews

Provocative and rewarding genetics

William J. Schull

Genetics, Evolution and Man. By W. F. Bodmer and L. L. Cavalli-Sforza. Pp. xv+782. (Freeman: San Francisco, 1976.) $13.95.

FIVE years ago, these authors gave us *The Genetics of Human Populations* (1971), a book addressed more to the specialist or advanced student than the neophyte. *Genetics, Evolution and Man* projects the notions set forth in this earlier effort on to a wider, less specialised readership. Their newer book expands the admirable features of its predecessor—excellent illustrations, conceptually and graphically; succinct abstracts at the beginning of each chapter, providing a sense of direction for the reader; the intercalation of frequent one-or two-line statements which capture the essence of the paragraphs which follow; problems which test and supplement one's understanding of each chapter; and a lucid, readily readable literary style. More important, perhaps, is its currency and the overall objectivity which Bodmer and Cavalli-Sforza bring to the task which they set for themselves. This does not imply that every human biologist will subscribe fully to their arguments on controversial issues—for example, the roles of heredity and environment in the observed differences among the races in tests of intelligence. Even where honest differences of opinion exist, however, it should be clear to all save the most committed that they make a serious appeal to reason rather than exhibit a preconception or prejudice.

This is unquestionably an important book, and everyone familiar with the subject matter must appreciate the wealth of information to which the un-initiated reader is exposed. It is provocative, but not provoking; it rewards without remonstrating. Each reader, indeed each reviewer, will undoubtedly be attracted to a different chapter. One of my favourites is the "Evolutionary Development of Modern Man". It attempts a synthesis of man's biological and social evolution as a human population geneticist might see it. The observations on which this synthesis is based are, of course, sensitive to substantial differences of interpretation both as to time and significance. Illustrative of such differences are the re-

ferences to the Jomon culture of Japan. Development of pottery by this culture is cited to be 9,000 yr ago, but *Japan; Its Land, People and Culture* (UNESCO 1958) suggests this event actually occurred about 14,000 yr ago; whereas the 'earliest Jomon' is given by Komatsu (*The Japanese People*, 1962) as about 6,400 yr ago. Apparently, Bodmer and Cavalli base their estimate on the carbon dating, carried out at the University of Michigan some years ago, of charred wood recovered from the Natsushima shell-mound in Kanagawa Prefecture. This specific example is not particularly important and quite possibly the figure given by Bodmer and Cavalli may be correct, but some of the 'facts' of their general argument will undoubtedly prove to be in error. This, however, is a much lesser moment than the synthesis and its respect for the divergent data sets which must, in some manner, be integrated. Their view does not disregard man's cultural evolution nor does it consciously serve the biological sciences. They acknowledge man's biological and cultural heritage; and seek

a quantitative appraisal of the contributions of these elements to his present diversity, not a capitulation on the part of one or the other disciplinary views.

If fault one must find, then this book is overly long. This stems, I believe, first, from a misdirected effort to make the book self-contained—that is, to provide the reader with every detail necessary to understand some of the subsequent arguments; and second, a failure to realise the frequency with which most students are exposed to the structure of DNA, the mechanics of mitosis, and elementary notions of statistics. This is not to suggest that reinforcement is not needed for some —possibly a large number—but a general application of repetition can just as readily lead to boredom as understanding. This is a small caveat, and doesn't seriously compromise the book's general excellence. ☐

William Schull is Director of the Center for Demographic and Population Genetics at the University of Texas Health Science Center, Houston, Texas.

Introduction to methods of geophysics

Geophysical Methods in Geology. (Methods in Geochemistry and Geophysics, Vol. 12.) By P. V. Sharma. Pp. xv+428. (Elsevier Scientific: Amsterdam, Oxford and New York, 1976.) Dfl. 65; $26.25.

HERE is a book that attempts to cover the techniques used in the whole field of geophysical research on the crust and lithosphere, industrial and academic; that is, a book that combines the contents of books such as Parasnis (1972), Bott (1971), York and Farquhar (1972) and Cox (ed. 1973). There are chapters on seismic methods (with sections on fundamentals of seismic wave propagation, seismic seismicity and seismotectonics, seismic prospecting, use of surface waves in crustal studies), gravity methods, magnetic methods (from Wenner spreads to magnetotellurics), radiometric methods (dating and prospecting), geothermal methods

and, finally, geophysics applied to global tectonics. The book is polymath, up-to-date and non-mathematical; SI units are used and a point in its favour is that most subsections end with a reference to a recent review article. It is a long book but not a verbose one.

Dr Sharma writes that the book developed from a course organised at the University of Copenhagen. "The course was aimed at providing a concise but fairly comprehensive introduction to the methods of geophysics within which the undergraduate students of geology and geophysics could locate their specific interests for further specialisation." He does not tell us in which year of their undergraduate career these students received this course. No man could have a lively research acquaintance with one half of the subjects touched on here so that the course (and the book) must have represented a con-

siderable piece of scholarship; but I feel that, in consequence, the book is at once too comprehensive and too facile to arouse the lively interests of an undergraduate reader. I shall continue to encourage the second- and third-year undergraduates whom I encounter to read Parasnis, Bott and Cox.   **D. Matthews**

Dr Matthews is a reader in the Department of Geodesy and Geophysics at the University of Cambridge, UK.

Planetary interaction

Interplanetary Encounters: Close-Range Gravitational Interactions. (Developments in Solar System and Space Science, 2.). By Ernst J. Öpik. Pp. vii + 155. (Elsevier Scientific: Amsterdam, Oxford and New York, 1976.) $26.95; Dfl.67.

THIS is an excellent book, a book that any scientist will read with awe, realising that very few of their number can, in the fullness of years, produce such a resumé of original work.

Ernst Öpik has been an astronomer of international renown since the early 1920s, his main interest centring on the Solar System. In this book he deals with the encounters between planets and between planets and stray bodies such as meteorites, asteroids and comets. These gravitational interactions sometimes lead to collisions— more often to changes in the orbital parameters of the two bodies and the possible ejection of one from the Solar System. Öpik has advanced from the restricted three-body problem to formulate a probability method reliant on two-body interactions over specific spheres of influence. This leads to a statistical celestial mechanics which culminates in the prediction of the cratering densities of Mars and Moon, dynamical lifetimes of cometary debri and other stray bodies, solutions to the Neptune–Pluto problem, stability of the Trojan asteroids, and many more phenomena.

The book summarises and refines a series of some 29 research papers written by the author between 1951 and 1972; it is beautifully produced, very clearly indexed and is essential reading for all Solar System astrophysicists.   **David W. Hughes**

David Hughes is a lecturer in the Department of Physics at the University of Sheffield, UK.

Quite recently, biological motivation has contributed new ideas to computer science and mathematics, while these disciplines have provided useful insights to biological development. North-Holland takes pleasure in detailing below three authoritative publications covering this new field.

Developmental Systems and Languages

by GABOR T. HERMAN *and* GRZEGORZ ROZENBERG.

With an introductory chapter by ARISTID LINDENMAYER, *Subfaculty of Biology, University of Utrecht.*

1975 xvi + 366 pages 10 tables 59 illus. over 200 lit. refs.
Price: US $28.95 / Dfl. 75.00; ISBN 0-7204-2806-8

The application of the rigorous techniques of formal language theory to biologically motivated concepts gave rise to the theory of developmental systems and languages. This area has been actively investigated over the last few years, yielding many results of fundamental interest to both the formal language theorist, the biologist and the computer scientist.

Self-contained, this book provides an exhaustive and up-date survey of the field. After an introduction in which the biological motivation of the study is described, the book is divided into three main sections. Part I gives a formal presentation of the theory of *developmental languages* while Part II treats the theory of *developmental sequences*. Part III offers a less formal discussion of topics whose biological motivation is more apparent.

Automata, Languages, Development

At the crossroads of biology, mathematics and computer science

edited by ARISTID LINDENMAYER *and* GRZEGORZ ROZENBERG.

1976 viii + 529 pages Price: US $46.00 / Dfl. 120.00; ISBN 0-7204-0474-6

A truly interdisciplinary exposition, this book represents the work of many distinguished scientists, brought together by their common interest in understanding the *algorithmic nature of developmental processes.* It includes both survey articles and specialized research papers covering recent trends in each of the following interrelated areas:

- mathematical and computer models for growth and development of multi-cellular organisms,
- the theory of L systems, and its relation to the formalization of multi-cellular development and regeneration,
- the theory of cellular automata and tessellation structures, and its relation to the concept of self-reproduction,
- the theory and applications of graph generating and related systems.

This book is based largely on the extension of material presented at a conference held at Noordwijkerhout, Holland, in April, 1975. Biologists, mathematicians and computer scientists will find this work a thoroughly representative and balanced review of the latest developments in an increasingly important interdisciplinary area.

Analysis and Control of Immobilized Enzyme Systems

edited by DANIEL THOMAS *and* JEAN-PIERRE KERNEVEZ.

1976 vii + 306 pages 22 tables 119 illus.
Price: US $27.50 / Dfl. 70.00; ISBN 0-7204-0361-8

This volume contains the proceedings of an international symposium held in Paris, May 1975 which was devoted to analysis and regulation of immobilized enzyme systems. The aim of the symposium was to bring together researchers from different fields, viz., biochemistry, applied mathematics and computer sciences, with the main emphasis laid on experimental and theoretical analysis of immobilized enzyme systems as biological models. The indentification of kinetic parameters as well as the control of enzyme systems ruled by partial differential equations, are presented.

obituary

Walter Schottky, author of many significant contributions to the fields of electron and solid-state physics, died on March 4, 1976 at the age of 89. Born in Zürich, he studied under Max Planck at the University of Berlin, and subsequently held university appointments in Wurzburg and Rostock. His first research interest lay in the field of electron physics, and what is now universally called the 'Schottky effect' (the increase in the thermionic emission of electrons from a solid resulting from an external electric field) dates from as far back as 1914.

Possibly as a result of the First World War, Schottky next turned his attention to the radio and in 1915 invented the screened-grid vacuum tube. He is also credited with the invention of the superheterodyne receiver. He studied the 'shot noise' which arises from the discrete nature of the charge carriers in a current of electrons, and in 1918 derived the well-known relationship between the magnitude of the noise and the strength of the current. For these three contributions he was awarded the Hughes Medal of the Royal Society in 1936.

After the war, Schottky's interests turned to thermodynamics and statistical mechanics. One result of this was his prediction of the 'Schottky specific heat anomaly'—the specific heat of a system possessing two closely spaced, discrete energy levels shows a maximum at a temperature comparable with the separation between the levels divided by Boltzmann's constant. His analysis successfully explains specific heat behaviour which has been observed on numerous occasions by low-temperature physicists. His postulation of the existence of 'Schottky defects' in crystals (missing atoms without compensating atoms in interstitial positions) also dates from this period.

In 1927, Schottky became closely associated with the firm of Siemens, thereby continuing a well-established German tradition. From that time onwards he turned his attention to semiconductors, and in particular to the properties of metal–semiconductor contacts, and it is for this work that he is probably most widely known today. It had been known since the early work of Braun in 1874 that metal points in contact with certain solids, such as lead sulphide, show rectifying properties, and although these point-contact rectifiers were not understood, they nevertheless played a most important role as detectors in the early days of radio. In 1931 Schottky and his co-workers showed that in copper oxide rectifiers nearly all the potential drop occurred near the metal contact, thereby implying the existence of some sort of potential barrier. By this time quantum mechanics was firmly established, and Wilson and others tried to explain the rectifying action in terms of the quantum-mechanical tunnelling of electrons through the barrier, but it was soon recognised that this mechanism predicted the wrong direction of easy current flow. Schottky, and independently Mott, suggested that the rectifying action resulted from the thermal excitation of electrons over the barrier, and showed that this model predicted the correct direction of rectification. Schottky supposed that the barrier region contained a space charge caused by a uniform density of charged impurities, so that the electric field increased as the metal was approached, whereas Mott assumed that the conditions of fabrication were such that the barrier region was devoid of impurities, so that the electric field was constant. We now know that Schottky's assumption is the more realistic one in practice, and in consequence the term 'Schottky barrier' is used almost universally to describe a metal-semiconductor contact. Schottky also derived an expression for the capacitance of the contact in terms of the density of charged impurities, and showed how measurements of the capacitance could be used to infer the impurity concentration. This technique is used very extensively throughout the semiconductor industry today.

Schottky was a very versatile physicist who combined a wide theoretical knowledge with a deep physical insight and the ability to look at problems from a practical point of view. He was always ready to apply his science to engineering problems, and it is perhaps appropriate that the term 'Schottky diode' is now firmly established in the vocabulary of electronic engineers who probably know little of his contributions to basic physics.

E. H. Rhoderick

The sudden death of **Professor Richard Foster Flint** of Yale University on June 6, 1976, deprived the world of one of the most eminent of its Quaternary geologists of any generation. Born in Chicago, Illinois, on March 1, 1902, the son of Professors Nott William and Edith Burnham (Foster) Flint of the University of Chicago, he obtained his B.S. in 1922 and his doctorate, *summa cum laude*, from that University in 1925. He joined the Yale University faculty as Instructor the same year and taught at Yale for the next 45 years. He retired in 1970 as Henry Barnard Davis Professor of Geology, the occasion being commemorated by a symposium in his honour and a resulting Festschrift (*The Late Cenozoic Ice Ages*) to which some of the world's leading researchers contributed.

Professor Flint's publications number over 150 research papers and a series of monographs including *Glacial Geology and the Pleistocene Epoch* (1947), *Glacial and Pleistocene Geology* (1957), and *Quaternary and Glacial Geology* (1971). The last stands as the most up-to-date, comprehensive, and masterful overview of the subject in any language. He was also the co-author of a number of editions of such famous Yale texts as *Physical Geology* and *Introduction to Physical Geology*. Especially influential was his stimulating teaching that opened the exciting vistas he saw to generations of Yale graduate students and undergraduates— thousands of the latter benefited from his introductory course on the earth sciences. He also enhanced Yale's reputation by serving as associate-editor, co-editor, or member of the editorial board of a number of scientific journals, including *American Journal of Science, Quaternaria, Quaternary Research, Radiocarbon* and *Zeitschrift für Geomorphologie*. In addition he was Chairman of the National Research Council committees that compiled the Glacial Map of North America and the Glacial Map of the United States.

Professor Flint once said that he did not expect great fame as a scientist, since he desired a life whose breadth would necessarily detract from single-minded scholarship. Yet he not only achieved the full life he sought, including a deep satisfaction in art and pottery, but contrary to his expectations, he also gained the world's acclaim as a scientist. He was a Fellow of the American Academy of Arts and Sciences, honorary member of many foreign geological societies and President of the 7th Congress of the International Quaternary Association.

Professor Flint leaves his devoted lifelong partner, Margaret C. H. Flint, a daughter, Anne Ogilvy, three grandchildren, one great grandchild, and a host of admirers, friends, and colleagues the world over.

A. L. Washburn

announcements

Appointment

Dr J. Ronayne to the Chair of History and Philosophy of Science at the University of New South Wales.

Professor Charles Antony Richard Hoare as Professor of Computation at the University of Oxford.

Dr Theodore Morris Sugden, Master of Trinity Hall, Cambridge, will be made President Elect of the Chemical Society in April 1977.

Professor Edward A. Smuckler as Chairman of the Department of Pathology, at the School of Medicine, UC-San Francisco.

Awards

The Paul Ehrlich awards for cell research have been given to **Professor John Gurdon,** of the Institute for Molecular Biology, and **Professor Torbjoern Caspersson,** of the Kungliga Karolinska Mediko-Kirurgiska Institutet, Stockholm.

The Royal Society Mullard Award for 1976 has been given to **Dr G. H. Hitchings,** recently retired from Burroughs Wellcome Co., for his work on chemotherapy, particularly the development of the drugs trimethoprim, allopurinol and azathioprine.

The Royal Society Esso Award for the Conservation of Energy has been made to **Mr. T. B. Jackson** of Honeywell Ltd.

The Ramsay Memorial Fellowships Trustees have awarded their General Fellowship to **Dr M. W. Evans** of the University College of Wales, Aberystwyth and the Netherlands Fellowship to **Dr M. R. Egmond** of the University of Oxford.

Meetings

September 22–24, **Systems and Models in Air and Water Pollution,** London (The Institute of Measurement and Control, 20 Peel Street, London W8 7PD, UK).

September 28–30, **Cellular Aspects of Neoplasma,** London (Dr Gisele M. Hodges, Department of Cellular Pathology, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK).

September 29, **Radioisotopes in Cardiology,** London (The General Secretary, British Institute of Radiology, 32 Welbeck Street, London W1M 7PG).

October 7, **Biology of the Pycnogonida,** London (The Executive Secretary, The Linnean Society of London, Burlington House, Piccadilly, London W1V 0LQ, UK).

Person to Person

Conoco Ltd announces the setting up of The Conoco Lecture on the Responsibility of Industry towards the Environment. The Conoco Lecture, which will become an annual event, will be presented to an invited audience in the spring of 1977 and will be submitted to the appropriate journal in the usual way. The Lecturer will have about four months to prepare his work and will receive a research award of £1,500. Nominations must be received by September 30, 1976.

Wanted, three- or four-bedroom flat or house, preferably furnished, in the West London area near to Ducane Rd. Will exchange for a three-bedroom flat in Turku, Finland from about October 25, 1976 to March 25, 1977. Replies (by September 20, please) to Docent Lauri J. Pelliniemi, M.D., c/o Shirley J. Green, Department of Histo-chemistry, Hammersmith Hospital, London W12.

Will send, for postage, small quantity of **Sophora chrysophylla** seed to persons interested in tropical highlands. Leguminous tree, frost tolerant, good timber, to 10 m high and 50 cm diam, Hawaii native. Perhaps of value in regions with climates similar to Quito, Ecuador. Apply to: W. Cook, PO Box 22364 Honolulu, HI. 96822.

Anyone going to Moscow who would like further information on the Moscow Seminars (see *Nature,* August 19) should contact: Joan Dale, Medical and Scientific Committee, 2 Frognal Rise, London, NW3.

There will be no charge for this service. Send items (not more than 60 words) to Martin Goldman at the London office. The section will include exchanges of accommodation, personal announcements and scientific queries. We reserve the right to decline material submitted. No commercial transactions.

October 14–19, **Warm Water Zooplankton,** Goa, India (Dr S. Z. Qasim, Director, National Institute of Oceanography, Dona Paula, PO Caranzalem, Panaji, Goa, India).

November 10, **Gerontology,** London (The Hon. Secretary, British Council

for Ageing, c/o The National Corporation for the Care of Old People, Nuffield Lodge, Regent's Park, London NW1 4RS).

September 5–9, 1977, **24th International Field Emission Symposium,** Oxford (Dr G. D. W. Smith, Dept of Metallurgy and Science of Materials Oxford University, Parks Road, Oxford, UK).

December 12–14, 1977, **Monitoring Hazardous Gases in the Working Environment,** London (Deadline for abstracts: January 31) (Dr John F. Gibson, The Chemical Society, Burlington House, London W1V 0BN, UK).

March 29–April 1, 1978, **Conference on Sub-Millimetre Waves,** Guildford, UK (The Meetings Officer, The Institute of Physics, 47 Belgrave Square, London SW1 8QX, UK).

August 16–23, 1978, **Plant Pathology,** Munich (Congress Plant Pathology, Biologisches Bundesanstalt, Messeweg 11/12, D 3300 Braunschweig, FRG).

August 30–September 6, 1978, **Virology,** The Hague (Fourth International Congress for Virology, Netherlands Congress Centre, PO Box 9000, The Hague, The Netherlands).

Reports and Publications

Philosophical Transactions of the Royal Society of London. A: Mathematical and Physical Sciences. Vol. 281, No. 1302: New Tensor Spherical Harmonics, for Application to the Partial Differential Equations of Mathematical Physics. By R. W. James. Pp. 195–221. UK £1.45; Overseas £1.50. Vol. 281, No. 1303: Continental Displacement and Expansion of the Earth During the Mesozoic and Cenozoic. By H. G. Owen. Pp. 223–291. UK £4.10; Overseas £4.25. (London: The Royal Society, 1976.) [14

The Zoological Society of London. Annual Report, 1975. Pp. 50. (London: The Zoological Society of London, 1976.) [24

Malaysian Rubber Producers' Research Association. MRPRA Publications, 1938–1974: An Author Bibliography. Pp. 106. (Hertford: MRPRA, 1975.) £2. [54

UKAEA, Harwell. AERE–R 8267: Radioactive Fallout in Air and Rain—Results to the End of 1975. By R. S. Cambray, Miss E. M. R. Fisher, J. D. Eakins and D. H. Peirson. Pp. 49. (London: HMSO, 1976.) £1.50 net. [54

Science Research Council. Daresbury 1975. Pp. 52. (Daresbury, Warrington: Librarian, Daresbury Laboratory, 1976.) [54

Australian Journal of Ecology, Vol. 1, No. 1, January 1976. Edited by Derek Anderson. Pp. 1–66. Published quarterly. Annual subscription £15; $A30 (Australia); $52.50 (N. America); £18 (overseas). (Oxford and London: Blackwell Scientific Publications, 1976. Published for the Ecological Society of Australia.) [54

Science Research Council: Engineering Board. Engineering Computing Requirements—Technical Group Report. (London: Science Research Council, State House, High Holborn, 1976.) *gratis.* [64

The Gale's Book of Countryside Projects. Pp. 24. (Carrow, Norwich: Countryside Offers, Colman Foods, 1976.) 10p. [74

Science and Technology in Islam. Pp. 48. (An exhibition at the Science Museum, London, 7 April—29 August, 1976.) (London: Science Museum, 1976.) [74

A Challenge for Change in the Dental Services. (A Consultative Document prepared by a Working Party under the Chairmanship of Mr. Laurie Pavitt, M.P.) Pp. 36. (London: Literature Sales, The Labour Party, Transport House, Smith Square, 1976.) 50p. [84

Tate and Lyle. Group Research and Development Annual Report 1975. Pp. 40. (Whiteknights, Reading: Tate and Lyle, Group Research and Development, 1976.) [94

Dream to reality in 25 years?

THE photocopier could hardly be called an unmixed blessing of 20th-century civilisation, but few would deny the substantial benefits it has conferred—or the number of jobs it has rendered unnecessary. Aware of the substantial cost (in foreign exchange) of photocopiers, the Indians set about designing one which could be produced by their domestic industry. In terms of the quality of copy delivered, the prototype devised at India's National Physical Laboratory in New Delhi, was as good as any model produced elsewhere. But, the innocent Western visitor asked, why is the loading of paper, the mounting of the material to be copied, the passage of exposed paper through developer still done by operators? Surely if you can apply high technology to the reproduction process, you can apply it to all other things going on in the machine. In our office the secretary just presses a button.

The reply to this rather insensitive question epitomises the dilemma of science and technology in the developing world—to make machines too automatic is to put people out of work, and the unemployed cannot look to a generous social welfare system to tide them over. Better a man employed on a modest wage doing a monotonous job; after all there is no shortage of manpower.

Such machines *versus* men, productivity *versus* employment issues have, of course, been raising themselves in a variety of guises all over the world for nearly two centuries, but this instance neatly illustrates the sort of clash in values that has endlessly to be resolved when developed technology encounters the developing world. These thoughts were occasioned by an important and little-reported lecture given, last week, by the Commonwealth Secretary-General, Mr Shridath S. Ramphal, former Attorney-General of Guyana, to the Science Policy Foundation in London.

Mr Ramphal's theme is that policymakers should be asking "what kind of science policy will contribute most to the eradication of poverty through a process of self-reliant economic advance that is consistent with social justice, environmental harmony and popular participation?" It is "appropriate technology" that should be at the centre of the development debate, not technology transfer. Technology, claims Mr Ramphal, is like genetic material in that it bears the code of social values of the society in which it was produced and sustained—export the technology to a region where the social values are different and it may even prove counterproductive.

A science policy in the developing world that is closely allied to, indeed subservient to, social policy, is not without its problems, as Mr Ramphal recognises. There are bound to be many scientists whose interests could never be channelled in socially useful directions, and yet it would be disastrous to start a witch-hunt to hound all genuine intellectuals out and into the first job available in the developed world. The problem here is in differentiating between the distinguished thinker whose very presence at a particular university or laboratory raises the overall quality of the work done and students produced there, and the hanger-on who attempts Western-style research incoherently and inconsequentially (Mr Ramphal's words).

Another major difficulty is the small size of many developing countries. Sixty of them have populations of less than 5 million, so the rapid accumulation of a "minimum critical mass of scientific talent" committed to well defined social values is not going to be easy. Ideally, regional centres and policies could be evolved, but the scientist's famed supranational spirit often means in reality regular trips half way round the world rather than just over the border.

Nor, Mr Ramphal points out, can the developed world stand back and let the developing world try and work it out alone—"the internal values of the world scientific community cry out for change . . . development [should] be the substantive goal of the science policy of the industrial world".

Many will believe that the case made, while strong on idealism, is weak on realism. The scientific community, pushed into being a tool of social policy, might react in a thoroughly conservative way amid cries about pursuing excellence for its own sake and truth wherever it may lead. And yet evolution is occurring, particularly among young scientists, in their views on the functions of science. Provided the pace does not carry things too far towards revolution, and provided that ideologists are not allowed to make all the running, thereby deterring many excellent but apolitical scientists, there is a real chance that what today seems like an unlikely dream for the developing world could, by the year 2000 be an effective reality. The issue is worth wider discussion in the developed world than it has so far received. □

Silk purse for sows' ears?

Managing Director **W. Makinson** exposes the workings, triumphs, and failures of the UK National Research Development Corporation (NRDC)

DESPITE the occasional protestation to the contrary the university inventor is accorded an important and respected status among the NRDC's clients. Without his contributions it is doubtful whether Sir Stafford Cripps' concept of a statutory body for promoting technical innovation would have survived the political events of the late 1960s and early 1970s. During the corporation's life, however, the attitudes of central government and the universities to research and its industrial application have changed markedly. The declining status of Britain in international trade, and, in particular, the failure of its manufacturing industry to compete with those in other developed countries has underlined the need to promote and utilise university research. The creation of wealth is arguably as important as the pursuit of knowledge, even at the expense of traditional academic freedoms. The NRDC's task, as defined in the 1948 Development of Inventions Act, is to assist in the transfer of new technology from the laboratory to industry, and because the corporation aims to be self-financing it has adopted arrangements that provide wherever possible for the recovery of the costs incurred.

It is hardly surprising, therefore, that the role of the NRDC in universities has come in for some critical examination. Those affected by new constraints, and those strongly motivated by the wish for closer and more productive relationships between academic and industrial communities, retain a deep interest in the directions followed by the organisation. This reappraisal has led to suggestions that technology emerging from universities might be better exploited by other routes than those offered by the NRDC itself.

There could be grounds for this new attitude. Some inventors are industrially oriented and have enough familiarity with legal, patent and commercial matters to go it alone. But they are rare. Experience shows that there are proven advantages in adopting a professional, centralised approach to exploitation, particularly where new technology is of international significance. Successful exploitation requires a careful blend of talents and resources: professional skills, a wide knowledge of industry, financial resources, an ability to assess and take risks and, on occasion, some ruthlessness in decision

taking are all essential. This is often not appreciated by academic research workers, to whom the invention is the major event rather than the starting point of a process fraught with problems and pitfalls, many of which he is neither qualified nor equipped to meet alone.

Moreover, experience also indicates that the majority of inventions turn out to have little commercial potential, however professionally they may be handled. And, paradoxically, it is the almost quixotic willingness of the NRDC to devote time and money to university cases which probably do not justify support and consequently fail, that has led to criticism. This cannot be lightly brushed aside, however: the corporation needs to retain the confidence of the academic research community.

Patent protection

The NRDC plays a crucial role in both establishing and protecting industrial development that may arise out of basic research. The world of patents is now so complex and has so many ramifications that it cannot be effectively handled without the aid of the specialised professionals. Industry can, in some fields, get by without patent protection and can penetrate markets simply through dynamism, confidential know-how, and competitiveness. Where the element of innovation is high, however, companies usually insist on adequate patent protection for as long as possible after products become established. This is especially true in the case of, for example, pharmaceuticals.

The executive officers of the NRDC are well supported by commercial, legal and patent services on a scale which compares favourably with all but the largest multinational companies; its patent department, for instance, is among the largest in the United Kingdom. These expert advisory services are available to university inventors even if they are not already clients; the earlier an enquiry is made, however, the more useful any advice is likely to prove.

Once an inventor has placed himself in the hands of the NRDC he must accept certain limitations. He is less free to dictate the subsequent course of development and exploitation, while his share of any subsequent revenue is inevitably cut. This may conflict with a

drive for recognition through immediate publication; it may also preclude direct personal arrangements with individual companies, even though that could be the most obvious route to early exploitation. Existing or future patent rights are assigned to the NRDC, and equity is maintained through a revenue-sharing agreement which divides proceeds from licence royalties or other income between the corporation and the university concerned (or the inventor). Under current arrangements, the rights to inventions arising from projects funded by UK research councils have, in most cases, to be assigned to the NRDC anyway.

The time that the NRDC takes to decide whether support is worthwhile has occasionally been a source of frustration. But some delay is inevitable. Potential markets must be properly assessed, as must the potential patent strength. Regrettably, the more commercially attractive propositions are more likely to run up against established competition. But throughout, the corporation maintains close contact with the inventor.

Finance for development

Unlike foreign government agencies with similar responsibilities, the NRDC can provide funds not only for research programmes which may eventually prove of commercial significance, but can also support the further development of inventions already made. Indeed, if the support for further development were not available it is likely that many university inventions could not be effectively exploited. Bearing in mind the NRDC's reserve borrowing powers and currently favourable financial status, there is, within reason, no upper limit to funds available for any particular project, as long as it can be effectively deployed. The slogan adopted by the corporation in 1973 : "£1,000,000 available for university research", is no idle statement.

The corporation always seeks the most effective route to commercialisation, so it usually attempts to involve an appropriate industrial organisation from the outset, either as a potential licensee and/or as a joint source of development funds. Negotiations may be carried out either by the university or the company, or both. It is quite normal for an inventor to act as a consultant on mutually acceptable terms.

Failure to arouse industrial interest in a particular project might often indicate that attempts at exploitation should be abandoned. On the other hand, it could be simply that a lack of sufficient evidence (in the form of a working prototype, for example, or of credible test results) is the only reason why industry cannot be persuaded that

an innovation is of importance—this is the "predevelopment gap" referred to in the Richards Report to the Engineering Board of the UK Science Research Council. While such a gap can be bridged by NRDC funds, the corporation will not shell out extra support merely to keep inventors happy; nor will it support an unnecessary research team, however competent. When further work is carried out in a university at the NRDC's cost a portion of any royalties is usually set aside to offset these expenses.

NRDC contributions to company development costs are recovered separately through a levy on subsequent sales of products exploited. The levy is calculated to provide a return reasonably related to the risks involved. The terms have from time to time been criticised as harsh, particularly when applied to proposals submitted by embryo, campus-based companies set up to exploit the results of research projects undertaken in a particular university. The criticism should be viewed against the statistical record, however, which shows that in three out of four cases the NRDC fails to recover its investment.

The corporation does not provide grants for company support, nor does it undertake to provide general working capital, except where an identified invention is involved. Companies must in such cases seek other sources of finance and accept the terms and conditions normally associated with them. It is worth noting, however, that support by the NRDC does not inhibit companies from qualifying simultaneously for other types of government assistance, nor need NRDC funding rank as a contingent liability where bank overdrafts or term loans are concerned.

Track record

Since its inception the NRDC has received about 5,000 proposals from universities, and taken assignment in about 2.000 cases. This compares favourably with the overall national picture including all sources of support —34,000 proposals and 6,000 assignments. So far, of the 2,000 university inventions supported by the NRDC about 200 have emerged as revenue earners; over the past five years this is about one out of every five or six submissions. The cumulative income attributable to university research is about £15 million, of which about £8 million has been recouped by the NRDC revenue-sharing agreements.

At present, the NRDC is involved in about 90 development projects involving universities, representing a total investment of about £1.5 million, about £1 million of which is actually being spent with the inventors. The largest single current commitment is that with
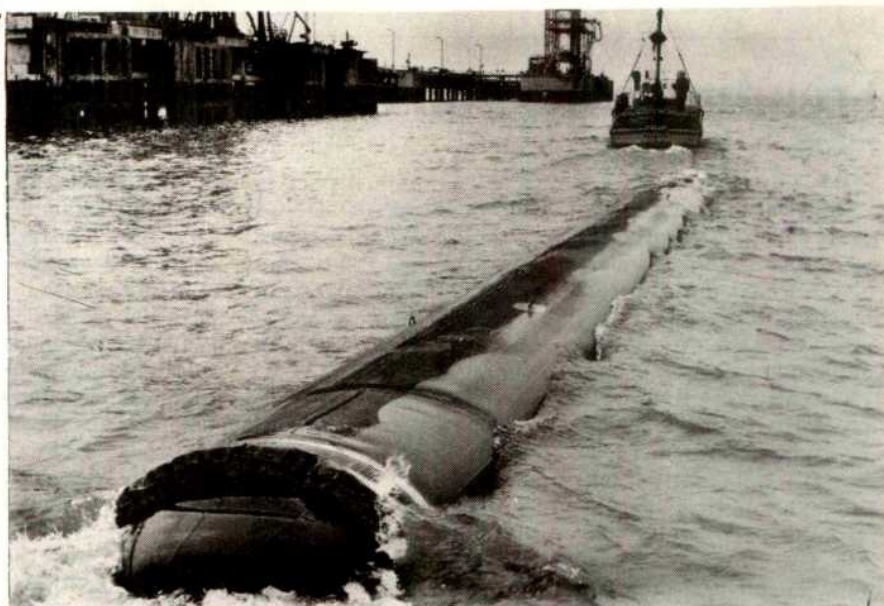
The Dracone, a soft failure

the Southampton Institute of Sound and Vibration Research where £160,000 is devoted to research into techniques to produce light-weight quiet, diesel engines.

By far the most conspicuously profitable inventions so far have been those made 20 years ago by research workers in the Sir William Dunn School of Pathology at Oxford in collaboration with the Antibiotics Research Station of the Medical Research Council. The work involved led to the discovery of the drug Cephalosporin C and the subsequent isolation and identification of its nucleus; worldwide licensing arrangements currently provide approximately 80% of the corporation's total annual royalty income. Other significant 'winners' would include the anticoagulant Arvin (Penang and Oxford), portable heart/lung machines (Royal Postgraduate Medical School), peptides derived from bee venom (University College, London) and methods of extracting diosgenin (the starting material for steroids used in the contraceptive pill) from fenugreek (Nottingham). Perhaps the most notable engineering inventions were those of Manchester University relating to computers.

More recent developments for which high hopes are sustained include improvements in electrochemical cells (Newcastle), continuous countercurrent ion-exchange techniques for uranium extraction (Imperial College, London), pioneering work on high modulus polymers, and novel flotation columns for mineral processing (Leeds), ultrasonics applied to metal-working (Aston), "surround" sound systems (Reading), speckle pattern comparators (Loughborough), and Pole Amplitude Modulation synchronous motors (Bristol). The NRDC has also set up a wholly-owned subsidiary company for the further development and marketing

of Genesys—a civil engineering computer-aided design system (CAD) initially developed at Loughborough University of Technology—and is in the process of creating a second software company for more generalised CAD packages. Though it is difficult to interpret statistics sensibly because of the often long gestation time of really successful inventions, the corporation is confident that the current input from university sources could maintain future royalties at a healthy level.

The NRDC has, of course, had its failures—sometimes spectacular—in its efforts to commercialise technically sucessful inventions. Any listing would have to include Dracones (large flexible floating containers for transporting fluids, based on proposals from Cambridge University) and the Tracked Hovercraft, which used the linear motor. Though the NRDC received no financial returns from either project both 'failed soft' in that the principles developed in these instances have been subsequently applied successfully to other industrial problems. Indeed, the Dracone, originally intended as an alternative to large oil tankers at the time of the Suez crisis, has been used to transport drinking water to the Greek islands, and might still play a useful role nearer home if the present drought persists.

Inevitably, many inventions from academic sources fail for simple technical reasons, usually in the translation from the original concept to practical realisation in commercial or industrial terms. Despite its years of experience, the corporation has no crystal ball; it will go on, no doubt producing more sows' ears than silk purses, but it has no intention of withholding support and encouragement wherever the rewards seem likely to justify effort and expense.   □

Treading softly on the ozone layer

A long-awaited report on the environmental effects of halocarbons was published in the United States last week. Colin Norman and Chris Sherwell report from Washington

THE US National Academy of Sciences (NAS) last week added its powerful voice to the scientific and political dispute about whether halocarbons, spewed into the atmosphere chiefly from aerosol spray cans, are causing serious damage to the Earth's ozone layer. The general press seemed to have some difficulty in deciding just what message the Academy was trying to convey, however. "Scientists Back New Aerosol Curbs to Protect Ozone in Atmosphere" ran the headline in the *New York Times*. "Aerosol Ban Opposed by Science Unit", said the *Washington Post*. Both were in fact correct, and the contrast reflects the finely balanced nature of the Academy's conclusions and recommendations.

The chief conclusion, set out in a long-awaited report by the NAS Committee on Impacts of Stratospheric Change, is that significant deterioration of the ozone layer will eventually occur if halocarbons continue to be released at their present rate. That would allow more ultraviolet radiation to reach the Earth's surface which, in turn would cause additional cases of skin cancer and other biological damage. Moreover, the committee warned of a possible effect on the world's climate from any continuing build-up of halocarbons in the atmosphere.

Alarming though such prospects may be, the committee cautions against taking precipitate action, such as banning the use of halocarbons as propellants in aerosol products. Instead, although it acknowledges that regulation is "almost certain to be necessary at some time and to some degree of completeness", the committee recommends that key uncertainties in the calculations should be cleared up before uses of halocarbons are restricted. The necessary information should be gathered in less than two years, the committee argues, and such a delay will cause little additional damage to the ozone layer.

The NAS recommendations were greeted with restrained enthusiasm by the halocarbon industry, and by the makers of hair sprays, deodorants and similar cosmetic products which together account for the major uses of the material. The industry-sponsored Council on Atmospheric Sciences said in a statement last week, for example, that the report backs the industry's view that "more research is required before any national decision on the fluorocarbon issue can be reached". Some scientists and environmentalist groups were not so happy, however. A spokesman for the Natural Resources Defense Council noted, for example, that the NAS recommendation for a two-year delay in regulation "is a value judgment which scientists are no better equipped to make than anybody else". Clearly, publication of the report will not put an end to the controversy which has been raging over the ozone depletion theory for the past couple of years.

The theory was first advanced in 1974 by two scientists from the University of California, F. Sherwood Rowland and Mario J. Molina (*Nature,* **249**, 810; 1974). They argued, in short, that the very properties which make halocarbons so useful—their insolubility and inertness—also pose a problem : unlike most airborne pollutants, they are not washed out of the lower atmosphere by rain. They gradually drift up into the stratosphere, where they are broken down by sunlight, releasing free chlorine atoms which break down ozone molecules through a complex series of chain reactions.

If that sequence of events is actually taking place, the result would be a reduction in the concentration of ozone in the stratosphere, and a consequent increase in the amount of harsh ultraviolet radiation reaching the Earth's surface. The ultimate consequence would be an increase in the incidence of skin cancer, which is correlated with exposure to sunlight, and possible biological damage to plants and animals.

Publication of the theory ignited a scientific and political controversy as some scientists attacked it as being implausible, while others argued that it is frighteningly realistic. The implications for human health also led to suggestions that the use of aerosol sprays be curbed, and prompted a massive research effort to test the validity of the theory.

Then, earlier this year, yet another alarming theory was put forward, namely, that accumulation of halocarbons in the atmosphere will result in a so-called "greenhouse effect", increasing the Earth's temperature and perhaps altering its climate.

The NAS committee has thus been poring over a mass of information accumulated since Rowland and Molina published their theory, and its report has been delayed for nearly six months by the appearance of important and conflicting pieces of data. In the end, however, the committee concluded that the evidence produced so far supports the ozone depletion theory. According to the committee's Chairman, John Tukey, a statistician from Princeton University and Bell Laboratories, the mechanism is now "a relatively well understood process".

It should be noted at this point, however, that it is impossible to test the theory by direct measurements of ozone concentration in the stratosphere since available measuring techniques are too insensitive to detect relatively small, long-term variations in the ozone layer. The theory is therefore based on laboratory experiments and computer analyses, supplemented with direct measurements of the atmospheric concentration of various key links in the chain reaction, such as chlorine, hydrogen chloride and so on. The whole business is a little like building up a jigsaw puzzle from badly cut and incomplete pieces.

Nevertheless, the committee has concluded that, if halocarbons continue to be spewed into the atmosphere at the rate they were released in 1973, the ultimate effect would be to reduce the concentration of ozone in the stratosphere by about 7%. It will take several decades to reach that level of depletion, however, since once released into the atmosphere, halocarbons can take up to a century to reach the stratosphere to do their work.

The figure of about 7% is deceptively precise. In fact, Tukey was careful to point out last week that it may turn out to be as low as 2% or as high as 20%, the uncertainties being caused by the complexity of the reactions—as many as 30 or 40 may be involved—and by the incompleteness of the information so far accumulated. That's one reason why the committee recommended against an immediate ban on non-essential uses in aerosols and suggested that more data are required.

Even if all release of halocarbons were to cease immediately, the committee has calculated that the material already in the lower atmosphere would continue to rise into the stratosphere and increase the destruction of the ozone layer for at least a decade. The ozone layer would then recover very slowly, with only half of the reduction being restored in 50 years. As for the consequences of delaying regulation by two years, the committee has calculated that if the figure of 7% reduction is correct, "whether a halving in (halocarbon) use and release were to take place in 1977 or 1979 would alter the

ozone reduction at any later date by no more than 1/6 per cent". In other words, "costs of delay in decision are small, not more than a fraction of a percent change in ozone depletion for a couple of years' delay".

Some perspective on the uncertainties in these calculations can be gained from a brief look at the problems which the committee encountered in putting out its report. The report was originally scheduled for publication in April but just as it was being put together, some new information, which seemed to throw the whole basis of the ozone depletion theory out of the window, suddenly came to light. In short, the theory would fall apart if there is a natural process which ties up chlorine atoms before they can attack the ozone molecules in the stratosphere. It seemed for a few weeks that just such a mechanism had been discovered. The suggested mechanism was the formation of chlorine nitrate, a compound which may be relatively stable under stratospheric conditions from chlorine atoms liberated from halocarbons and oxides of nitrogen already in the stratosphere.

When calculated rates of formation of chlorine nitrate were slotted into the equations, it turned out that halocarbons may even have a positive effect on the ozone layer—in other words, they may increase ozone concentration by removing oxides of nitrogen from the stratosphere. Later computations indicated, however, that the assumed rate of formation of chlorine nitrate had been greatly overestimated, and the effect of the mechanism is actually quite small.

As for the potential climatic effects of halocarbons, the committee suggests that precise information is equally hard to come by. Nevertheless, its best guess is that there will be a warming effect due to absorption of infrared radiation by halocarbons in the atmosphere, the effect being similar to that proposed for carbon dioxide accumulated through the burning of fossil fuels. Tukey suggested that the effect from halocarbons in the year 2000 might be about 40% of that due to burning fossil fuels, but he added that the relationship between such warming trends and broader climate changes is very poorly understood. "There are going to be some climate effects and it would be, I think, a disservice to put hard numbers on them this year", he said.

Although the committee recommends against immediate restrictions on the use and release of halocarbons, it does urge a number of steps to prepare the way for regulation should it be needed. At present, the authority to regulate the use of aerosol sprays is vested in at least three different agencies—the Food and Drug Adminis-

tration (FDA) the Environmental Protection Agency (EPA) and the Consumer Product Safety Commission (CPSC)—and it is questionable whether any of them really has a powerful enough mandate to act. Legislation is clearly required. Last week, by coincidence, Congress finally approved the Toxic Substances Control Act, a legislative landmark which, according to some interpretations, may provide EPA with enforceable authority over halocarbons. But the Bill is opposed by industry and may be vetoed by President Ford. Similarly, Congress is close to passing a series of amendments to the Clean Air Act, which again would give EPA a powerful lever to regulate aerosol products. The problem here, however, is that the House and Senate may not be able to settle some basic disagreements on other parts of the Bill before the October 2 recess. Thus, legislation may have to wait until next year.

The committee does, however, recommend one potentially important step which could be implemented relatively quickly. It suggests that aerosol sprays containing halocarbon propellants should bear a label stating that fact. The FDA, which has authority to regulate cosmetics, probably has the power to enforce such labelling. Though the committee argues that

"labeling should in no sense be regarded as a substitute for regulation but rather as an aid to consumer self-restraint", the effect could be rather large, since consumers are already beginning to turn away from some aerosol products. In 1975, for example, the use of halocarbons declined by 15% —the first decline in more than a decade (the accompanying table shows 1975 figures).

As for the effects on industry of a ban, the committee says they would be "appreciable". Just how appreciable can be gleaned from the fact that total world production of the two halocarbons with which the committee is concerned—namely, $CFCl_3$ or F-11, and CF_2Cl_2 or F-12—amounted in 1974 to nearly one million tonnes. But, says the committee, the industrial consequences do "not loom large" against the background of "a possible, although very small, change in world climate".

Finally, it should be noted that the US accounts for almost half the total world use of halocarbons. Thus, if the US Government decides on regulation, it would only attack half of the problem. The committee therefore recommends that other countries should be encouraged "by whatever appropriate means are likely to be effective" to take similar action.   □

Estimated worldwide releases of F-11 and F-12 in 1975 (millions of pounds)

Aerosols **1115.1** (74.5%)	Personal	Antiperspirants/deodorants		458.4
		Hair care		401.5
		Medicinal		37.3
		Fragrances		2.3
		Shave lathers		0.9
		Others		34.4
	Household	Room deodorants		17.7
		Cleaners		9.6
		Laundry products		23.4
		Waxes and polishes		9.2
		Others		9.2
	Miscellaneous	Insecticides		33.3
		Coatings		22.9
		Industrial		39.0
		Automotive		8.0
		Vet. and Pet		2.3
		Others		5.7
Air conditioning **refrigeration** **204.7** (13.7%)		Mobile air conditioning		89.8
		Chillers		42.9
		Food store		33.1
		Beverage coolers		5.8
		Home refrigerators and freezers		5.8
		Miscellaneous		27.3
Plastic foams **176.5** (11.8%)		Open cell		100.0
		Closed cell		76.5
			Total	**1496.3**

Figures are based on the annual incremental releases of F-11 and F-12 indicated for 1975 in a 1976 report by the Manufacturing Chemists Association, and the detailed percentage analysis by users in the US for 1973 in a 1975 report by Arthur D. Little, Inc. for the EPA. Figures do not include USSR and Eastern Europe.

Breakdown between F-11 and F-12 not available. If F-11($CFCl_3$, 77% chlorine by weight) could be replaced by an equal weight of F-12 (CF_2Cl_2, 59% chlorine by weight)—not often feasible—the amount of ozone reduction would be decreased.

Source: *Halocarbons: Environmental Effects of Chlorofluoromethane Release*, Chapter 1 and Appendix D.

Recipe for less disaster

The first ever report of the recently formed UK Advisory Committee on Major Hazards (CMH), which recommends improved safety measures for potentially hazardous industrial installations, was published last week. Alastair Hay reports

THE new measures advocated in the CMH report* could provide a framework for the safer building and siting of industrial plants in Britain. The report considers the hazards of "explosion, the sudden release of toxic substances (and) cataclysmic fire", and industries that would come under closer scrutiny include those producing fertilisers, explosives, petrochemicals and nylon and other polymers. Nuclear installations are not considered in the report.

The report is the first of its kind to be produced as a result of a government commission in any country. The CMH was appointed by the Health and Safety Commission (HSC) following the disaster at the Nypro chemical plant at Flixborough in June 1974, which killed 28 people, injured more than 100, and resulted in an estimated £28 million worth of damage to the immediately surrounding area. The Flixborough explosion was the most devastating in Britain since the end of World War II.

Prominent among the report's 38 recommendations is a call for a notification scheme that would allow the UK Health and Safety Commission (HSC) to identify particularly hazardous installations. In drawing up a list of notifiable installations the committee identifies 8 different types of plant, but ignores nuclear installations, in line with its original brief. An estimated 1,000–3,000 sites at present operating within the UK would fall within the proposed list, of which at least one in ten is operating to safety standards giving possible cause for concern, the report suggests. Moreover, while potentially notifiable plants are distributed throughout the UK, the report identifies specific areas of concentrated development such as Teesside and Canvey Island where a chain reaction of explosions might pose an additional hazard. The proposed list of notifiable installations includes:

● those storing or processing toxic material where there could be an emission of toxic gases or vapours equivalent in effect to more than 10 tonnes of chlorine.

● those storing or processing flammable materials where there could be a rapid emission of more than 15 tonnes of flammable gases or vapours.
● Those storing or processing more than 5 tonnes of materials which are intrinsically unstable or of very high exothermic reactivity. Examples are ethylene oxide, acetylenes, and organic peroxides.
● Those with a large inventory of stored pressure energy, operating at 100 bars or more, using gas phase reactions.
● Those storing or processing more than 10,000 tonnes of flammable materials with a flashpoint of less than 22 8 °C.
● Those storing or processing more than 135 tonnes of liquid oxygen.
● Those storing or processing more than 5,000 tonnes of ammonium nitrate.
● Those storing or processing materials which during a fire could cause an emission of toxic gases or vapours equivalent in effect to more than 10 tonnes of chlorine.

The list would apply to both existing and proposed installations, though the CMH has stressed that it is only provisional, and that a future list will almost certainly be wider ranging.

The report underlines the increasing importance of getting design and operating procedures right the first time. "Because of their present-day size and throughput," the report says, "there are now many plants throughout the world where a critical first mistake can result in disaster". Moreover, the committee warns of the increased dangers to local communities: the report notes that while "the probability that an individual worker will be involved in a fatal accident has notably fallen, the chances [of] a plant failure . . . involving the public at large . . . [have] become considerably greater".

Some of the specific proposals included in the report stem directly from recommendations made by the court of inquiry into the Flixborough disaster. For instance, the report advocates that industrial pipework, valves and pumps, at present often only tested before assembly, should also be tested following installation, when they form part of a total pressure system. It also recommends that any changes in the operating procedure at plants should be reported to the HSC for evaluation. Control rooms—that at Flixborough, built of brick, was completely destroyed in the explosion—should be single storey, reinforced concrete buildings, situated in a relatively safe area, while black box recorders similar to those used in air-

craft, should be introduced to monitor plant operations, allowing the identification of fault sequences leading to accidents. Moreover, safety officers should have easy access to international data banks on industrial accidents and plant safety features, the report says.

The report advocates closer consultations between local authorities and the HSC when planning permission for new industrial plants is being considered. Neighbouring authorities likely to be affected by final decisions should also be included in the discussions. But the CMH has not yet considered the consequences of closing down potentially hazardous sites already in operation. At present, the compulsory closure of existing plant involves local authorities in the payment of considerable sums in compensation. The committee plans to discuss this problem with the Department of the Environment.

The Report does not advocate the enforced notification of 'near misses', which, it feels, could inhibit the flow of information. Mr James Tye, director of the British Safety Council, welcoming the report, estimated that for every fatal accident in British Industry, there are three involving injury to workers, thirty resulting in damage to plant, and six hundred 'near misses'. Nonetheless, both industry and the HSC recognise that the lessons to be learned from a 'near miss' are as valid as those gained in the wake of a full-scale accident.

Professor Bryan Harvey, Chairman of the CMH, stressed the responsibility of management in introducing appropriate safety measures even before they became compulsory. Industry, he said, would in future have to give far more consideration to safety factors. Speaking for the chemical industry, Mr Donald Bennett, Chairman of the Chemical Industry Safety and Health Council, welcomed the report as "broadly consistent with the industry's own thinking on the form of control required". But he singled out areas needing further investigation, among them "the definition of the unit of notification", what parts of plants, processes or sites should be licenced, and the problem of selecting "truly independent experts" to act as arbitrators between industry and the HSC.

Mr Bennet's sentiments are not inconsistent with those of the CMH itself, outlined in the introduction to the report. The committee acknowledges the need for further discussion and research, and admits to the preliminary nature of much of the work completed so far; the report has been distributed to various interested organisations with requests for specific comments on the committee's proposals. □

BRAZIL

CONCERN for the environment is increasing in Brazil, a country which until recently dismissed ecological preoccupations as detrimental to industrial development. But even though public consciousness about the environment in this vast nation of 110 million people has been raised, little has been achieved in practical terms, and the environmental battle is far from won.

On the plus side, however, a new Presidential-level, environmental protection office—known as SEMA—has been authorised to veto state financing of industrial undertakings lacking adequate pollution safeguards. The head of SEMA, Paulo Nogueira Neto, regards this as a major victory for the environmental protection movement. And he does not believe it will be very hard to persuade Brazilian industry to fight pollution before producing it "The inclusion of antipollution equipment in an industrial project usually adds only 2% to the project's total cost", Nogueira Neto declared recently. "On the other hand, if a factory is ordered to install antipollution equipment after it already has gone into operation, the cost of such equipment will be much higher." (There's a major catch to this, however : SEMA does not yet have adequate authority to shut down existing transgressors, and cannot force them to install antipollution gear.)

Also on the plus side, the Federal Government's much criticised and undermanned Forest Service has scored some recent successes in catching clandestine hunters, and in preventing the export of hides of illegally slaughtered protected species such as deer, monkeys and alligators. The Forest Service plans to burn 100,000 recently apprehended pelts and skins, hoping to deter future unlicensed hunting. Most illegal hunting takes place in Mato Grosso State and the Amazon Jungle, which, because of their immense size and lack of roads and communications, are practically impossible to patrol.

The Forest Service has also won a victory of sorts by successfully fining the Brazilian subsidiary of Volkswagenwerk AG the equivalent of $25,000. Volkswagen has gone into the cattle ranching business in the Amazon Jungle, taking advantage of liberal Brazilian tax incentives. The fine follows a technical violation concerning the clearing of trees on Volkswagen's property: The Forest Service at first wanted to fine Volkswagen the equivalent of more than $5 million for the alleged un-

authorised felling of some 9 million trees, but the plan was reportedly shot down at higher levels within the government, which wants to maintain a favourable climate for foreign investment.

On the minus side of the environmental picture, nearly 2,000 residents of a slum district in the north-eastern city of Salvador were hospitalised after breathing potentially deadly chlorine gas from a chemical plant. The plant is an indirect subsidiary of Petrobras, the state oil company, and will probably escape with only a fine.

Salvador is, in fact, apparently a major target of industrial polluters. The neighbouring Atlantic Ocean and several rivers in and near the city are becoming poisonous repositories of mercury, cadmium, sulphuric acid and ferrous sulphate. Jun Ui, a Japanese expert, has studied the situation in Salvador and concluded that the population is on the verge of mass poisoning, similar to the mercury poisoning that caused a scandal in Minamata, Japan, during the 1950s. Professor Ui has found traces of mercury in the hair of Salvador residents, and has said that cadmium in the region could cause bone disease.

People in several Brazilian cities have begun to band together to agitate for the preservation of urban green space; so far, however, none of these groups have been very effective. In Sao Leopoldo, a suburb of Porto Alegre, schoolchildren recently held a public demonstration against the felling of stately palm trees within the school grounds. The Forest Service fined the school administration and ordered it to stop, but not before 25 coconut palms had fallen to the chain saw. In Belo Horizonte, a growing city of 1.5 million people that is rapidly turning into a "concrete jungle", citizens meanwhile took to the streets to protest against the felling of a grove of trees on Church land.

But the Church obtained permission to continue the clearance, and now plans to build a shopping centre and office building on the newly stripped site.

● An antismoking lobby has emerged in Brazil. The Rio Grande do Sul State Medical Association has managed to scrape up the equivalent of $40,000 for a three-month advertising campaign warning of the health hazards of smoking. It aims to persuade the Brazilian news media to carry antismoking propaganda to offset cigarette advertising. There is also pressure for government health warnings on cigarette packages and tobacco advertisements. The antismokers face powerful adversaries. Brazil's cigarette industry, dominated by a subsidiary of the giant British-American Tobacco Company, is one of the biggest advertisers in the Brazilian news media, and pays more than $1,500 million in Brazilian taxes each year.

● Brazilian petroleum engineers have always believed that there is oil in Brazil's Amazon region, a belief not generally shared by the world's oil industries. Now, however, Petrobras has found oil off the coast of the far northern territory of Amapa, and along the continental shelf off Rio de Janeiro State. It is still too early to tell how much oil may be in the area of the test-well but the fact that Brazil has finally discovered Amazonian oil is important in itself. Brazil, whose principal domestic source of oil is the north-eastern state of Bahia, still imports 80% of its annual oil requirements, which causes serious balance-of-payments problems. Five international oil corporations plan limited prospecting operations in Brazil.

● A Brazilian obstetrician is encouraging his patients in the prosperous middle-class city of Curitiba, in the southern part of the country, to have their babies in the same way as the native Indians: squatting instead of lying down. Dr Moises Paciornik has studied primitive Indian tribes in Brazil and has concluded that squatting births lead to less pain and fewer postnatal complications. Once contractions have started, delivery is safer and quicker, if gravity is allowed to help, Dr Paciornik says. He has adapted one of the delivery rooms in his Curitiba clinic for squatting births, complete with an Indian-style mat on the floor, and he says his urban patients show surprisingly little reluctance about giving birth the Indian way.

Bruce Handler

IN BRIEF

Electric cars for US

Congress last week won a small, but important, skirmish with the Ford administration over a plan to develop an electric car industry in the US. It overrode a Presidential veto on a Bill to devote up to $160 million to research, development and demonstration of electric car technology over the next six years. The key part of the Bill—and the part which stuck in Mr Ford's throat—is a scheme to create a market for electric vehicles through direct government support. In the next three years, the Government is empowered to purchase up to 2,500 electric vehicles, with a further 5,000 within five years. Mr Ford had vetoed the legislation feeling that the technology was not sufficiently advanced for a subsidised purchasing programme of that size.

Engineers building up

For all that is being said about the sad state of recruitment into engineering, figures just released by the Universities Central Council on Admissions (UCCA) show a steady growth in the number of applicants, and admissions, to universities for engineering courses (Statistical Supplement to the Thirteenth Report, 1974–5, UCCA, £1.25). Since 1973 electrical civil and mechanical engineering have each registered gains in applicants—firm figures for 1975, and reliable forecasts for 1976 show that in three years engineering has gained 30% more applicants. Mathematics, physics and chemistry remain in the doldrums; chemistry was a few per cent up both in 1975 and 1976, but physics dropped several per cent in 1975 and just held its ground in 1976.

Soyuz in orbit

Carrying photographic equipment developed partly by the East German company Carl Zeiss (Jena), the latest Soviet space-shot, Soyuz 22, went into orbit last week. Special emphasis is being paid to the photography of large tracts of Soviet territory, and while practical results are anticipated—similar, it is hoped, to those from the Soyuz 12 survey which revealed potential oil- and gas-bearing zones in the region of the Mangyshlak and Buzachi peninsulas—the main aim is to develop new techniques. Additionally, as a follow-up to earlier experiments during which cultures of *Proteus* bacteria grown in space showed signs of developmental deviations, experiments aboard Soyuz 22 have been designed to investigate the phenomenon.

SCIENTIFIC meetings used to be almost routinely enlivened by projectionists who inserted slides upside down, backwards, and so on. But the introduction of carousel projectors has largely done away with this form of amusement, especially when the lecturer is artful enough to load the slides in advance. I always try to meet projectionists ahead of my talks. Most of them are males, often students who take the job as part-time employment. I usually scrutinise the projection booth for copies of *Sports Illustrated* and soft-core porn magazines, which I scan briefly to soothe my nerves. If present, such reading matter shows that the projectionist will not be listening for slide-changing cues during my lecture, and I know that I must shout my requests.

One of my adventures as a lecturer had an almost dream-like quality. It was in Las Vegas. I arrived at the "meeting hotel" to speak to a convention of dietitians. After threading my way past slot machines and blackjack tables, I found that the theme of the day was "Nutrition in the Age of Aquarius". The delegates wore emblems proclaiming their zodiacal signs. Some of them were multiparous married women whose badges stated "I am a Virgo", which seemed incongruous to me. The slide projector had somehow been overlooked, and it was one hour to launch time. A hotel employee was delegated to drive me on a hunt for the machine. We bowled along a road that headed west, past shacks, into the relentless desert landscape until I realized that we were well on the way to Death Valley. Eventually we backtracked and spotted a Quonset hut that yielded a pro-

jector, and we got back to the casino in time. I was so stimulated by the experience that I preceded my lecture with an impromptu summary of my unvarnished opinion of the zodiac and of horoscopes.

Previously in these pages, I have described the incident in which

---

Slide slips

THOMAS H. JUKES

Marianne Grunberg-Manago disappeared from sight at the start of her lecture, in a manner similar to the descent of Erda in *Das Rheingold*. I made a mistake; I have received a peremptory letter of reprimand from the head official of the American Society of Biological Chemists, pointing out that it happened in Philadelphia in 1958 rather than in Atlantic City in 1963. Another correspondent, more sympathetic, says that he re-

members that Marianne blew kisses and waved to the audience as the platform went down. Be that as it may, I recall another Federation Meeting—I think it was in Chicago—addressed by Larry Irving, in which some interesting slides were shown at a long lecture on respiratory physiology. They were remarkable for all being projected sideways. The large audience developed a transient form of torticollis resembling a sort of a lateral whiplash injury. I believe that the explanation for the misoriented slides had a nationalistic background. It seems that British slides used to be $3\frac{1}{4}''$ square, American slides $3\frac{1}{4}'' \times 4''$, while the Canadians, with their altogether admirable trait of self-assertiveness, used slides that were $4''$ square. The slides in question must have been prepared by a Canadian for showing in US projectors, which the Canadian evidently thought received $3\frac{1}{4}'' \times 4''$ slides vertically rather than horizontally.

One of the new amenities is an antenna that is coiled around the lecturer like the string around a parcel. I canot vouch for the following story that was related to me, but even if it didn't happen, it has obvious possibilties for users of Citizens Band radios. An unfortunate scientist swathed in his antenna, switched on the microphone and started his talk. Or tried to. But instead of the lecturer's voice the equipment picked up, on the same wavelength, the thunderings of a fire-and-brimstone preacher at a nearby revival meeting. I am told that the scientist plucked the offending antenna from his body in a vain effort to quell the interruption.

news and views

X-ray cryoenzymology

from C. C. F. Blake

ALTHOUGH the degree of knowledge constituting an "understanding" of the catalytic power of enzymes may be a matter for debate, the basic nature of certain kinds of information is not in dispute. These essential pieces of information include the three-dimensional structure of the enzyme, and especially its active site; the sequence of enzyme–substrate intermediates that occur along the reaction pathway, their individual structures, and the energetic relationships between them. Clearly no single technique can provide information at the level required on all these topics, and a synthesis of data from several disciplines is necessary before correct answers in a very difficult field can be expected.

The contribution of X-ray diffraction, as the unchallenged technique for the determination of three-dimensional structure, would seem to lie in principle in defining the structures of the native enzyme (with essential aid from sequence studies) and the enzyme–substrate intermediates. But, despite great success with the enzymes themselves, structures of true, productive intermediates have proved elusive. The reason for this is that at room (or physiological) temperatures there is such a gross inequality between the rapidity of the enzyme reaction even in the crystalline state, and the length of time required to complete an X-ray experiment, that there is no possibility of success. The only productive intermediates that can possibly be examined in normal conditions are the Michaelis complexes of certain enzymes whose reactions and equilibrium constants are favourable.

The inability to look at true enzyme–substrate intermediates has not prevented X-ray diffraction from producing some important results on the activity of certain enzymes, for example, lysozyme, the serine proteinases and carboxypeptidase. With these enzymes, results from the study of the interactions of inhibitors or virtual substrates could be extended by structural and chemical considerations to produce plausible models of the productive complexes, from which generally accepted hypotheses of their hydrolytic activity have been deduced. However, the intracellular metabolic enzymes that are now under study have so far resisted this approach, presumably because their more sophisticated mechanisms are not so amenable to chemical intuition. In this context the two papers in the current issue of *Nature* (Fink and Ahmed, page 294 and Alber, Petsko and Tsernoglou, page 297) on the formation and X-ray analysis of stable, productive enzyme–substrate intermediates at very low temperatures, represent a radical departure of great promise.

It has long been known that many enzymes are catalytically active at sub-zero temperatures. Douzou, in his pioneering studies of enzymes at low temperatures (Douzou *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **66**, 787; 1970), has shown that true enzyme–substrate intermediates can be isolated in a stable form at very low temperatures if sufficient care is taken with the physicochemical properties of the aqueous–organic solvent systems that are necessary. In view of this Fink and Ahmed have investigated the behaviour of chymotrypsin, trypsin and elastase towards specific ester substrates at temperatures between −20 and −70 °C in aqueous–organic solvents, to test the feasibility of obtaining stable crystalline intermediates of these enzymes suitable for X-ray studies. They show that in these conditions the rate of acylation is several orders of magnitude faster than the rate of deacylation, and that for each enzyme, conditions can be achieved in which the rate of deacylation becomes so low that acyl-enzyme intermediates accumulate to high concentrations. At temperatures of −50 to −70 °C in aqueous dimethyl sulphoxide or methanol at *p*H values of maximum activity, they have prepared well-defined acyl-enzyme intermediates in almost stoichiometric quantities, for all three enzymes both in the dissolved and crystalline states. These intermediates are stable for periods of days to months at low temperatures, but when warmed up to 0 °C they are turned over at the expected rate.

A preliminary X-ray analysis of one of the stable acyl-elastase intermediates characterised by Fink and Ahmed is reported by Alber, Petsko and Tsernoglou. Elastase was chosen because its active site is accessible, and it exhibits low solubility and high stability in the aqueous methanol which is necessary for the X-ray studies because its low viscosity allows the substrate to diffuse easily into the crystal. Using a flow-cell they were able to change the mother liquor from 0.01 M sodium acetate to aqueous methanol in a number of stages of increasing concentration and lower temperature until 70% methanol was reached at −55 °C. In these conditions the substrate, N-carbobenzoxy-alanyl-*p*-nitrophenol ester could be added to the cell and its binding followed crystallographically. After equilibrium was attained three-dimensional X-ray data were collected to 3.5 Å resolution. As the original structure determination (Watson, *et al.*, *Nature*, **225**, 806; 1970) was of the tosyl enzyme, Alber, Petsko and Tsernoglou independently determined the structure of the native enzyme at room temperature, and also collected native data at −55 °C. This enabled them to be clear that the enzyme's structure was essentially undisturbed by the transfer from an ionic liquid at room temperature to an organic liquid at low temperature. The difference map calculated from the data sets collected at −55 °C, although not analysed in any detail in accord with the modest resolution limit, was nevertheless fully consistent with the expected productive acyl-enzyme intermediate in the catalytic process as predicted by Fink and Ahmed. The control experiment of raising the temperature of the crystal to −10 °C caused the binding curve to reverse, and the corresponding difference map showed no density remaining in the active site.

It seems certain that this technique will in due course be used to produce

detailed structural information on the acyl-enzyme intermediate, and probably also on other intermediates, of the serine proteinases. This would come close to realising the full potential of X-ray diffraction studies of enzymes, that of being able to obtain a series of "stills" of the enzyme–substrate complex at each point of stability along the reaction pathway, which would indeed be a major contribution to the understanding of enzyme activity, however this understanding may be defined. ☐

Cytoplasmic control of protein synthesis

from Pamela Hamlyn

An EMBO workshop on Cytoplasmic Control of Eukaryotic Protein Synthesis was held on July 19-22, 1976 at King's College, Cambridge and was organised by Tim Hunt and Richard Jackson, Department of Biochemistry, University of Cambridge, and Alan E. Smith, Imperial Cancer Research Fund, London.

FOR many workers in the field of eukaryotic protein synthesis, control of initiation of polypeptide formation is synonymous with translational control. A suitable starting point for the conference was, therefore, to consider the protein initiation factors which regulate the various steps during initiation. Two groups have purified to homogeneity the initiation factors from reticulocytes. W. C. Merrick described the work of the National Institutes of Health group and T. Staehlin that of the Basel Institute for Immunology. Each group has its own system of naming the different initiation factors, but the physical and functional characterisation of the proteins is sufficent to allow them to make correlations with confidence. Four initiation proteins have been isolated from Krebs II ascites cells. H. H. Dahl (ICRF, London) explained their functional activities and described experiments in which reticulocyte factors had been successfully substituted for ascites factors. In the past some initiation factors were claimed to be messenger-specific. Most workers now think that such observations were due to inadequacies in the assay systems used and agree that initiation factors are functionally homologous in different tissues and species.

Tails and caps
Messenger RNAs may have poly(A) tails, 'caps' (blocked and methylated 5' termini) and untranslated sequences either side of the amino acid specifying region. Since these structures may be involved in translational control mechanisms, part of the conference was devoted to them. G. Marbaix (University of Brussels) confirmed that the poly(A) tail confers stability on globin

mRNA injected into oocytes and showed that messenger lacking poly(A) was only degraded once its translation had been initiated. R. Kaempfer (Hebrew University, Jerusalem) showed that globin mRNA requires the 3' non-coding region for its translation but not the poly(A) tail.

The work on the cap region is in an earlier stage, at which many experimental results seem contradictory. In his summary of this group of lectures the chairman, B. P. Perry (Institute for Cancer Research, Philadelphia), concluded that the effects of cap on translation were quantitative and seemed to depend on the messenger, the cell-free system used and the conditions of translation.

Inhibition
The section on the control of initiation centred on the nature of the inhibitor of protein synthesis formed when a reticulocyte lysate is incubated without the addition of haemin. Isolation of the inhibitor results in the co-purification of a phosphorylase activity which can phosphorylate the initiation factor responsible for the formation of the ternary complex between GTP and met-tRNAf. The phosphorylated factor is inactive. Does the phosphorylation cause the inactivation, or is it in some way coincidental? The group from the University of Cambridge have isolated an inhibitor which specifically phosphorylates *de novo* the small subunit of initiation factor IF-E2 (IF-MP). The copurifying kinase and inhibitory activity are heat inactivated at the same rate. Further evidence that phosphorylation of IF-E2 is the inhibitory event is provided by experiments of B. Hardesty (University of Texas) whose group has shown that antibodies to the inhibitory factor, which prevent it acting as an inhibitor, also prevent the phosphorylation of IF-E2. The group from MIT have broadly similar results but are more cautious in their interpretation, pointing out that phosphorylation of IF-E2 may be a normal event at some stage in protein synthesis and that the activation of the inhibitor may alter the normal phosphorylation pattern of IF-E2 and somehow prevent the recycling of the factor. The work

of M. J. Clemens (National Institute for Medical Research, London) indicates that phosphorylation alone is not sufficient to inhibit binding of IF-E2 to the GTP-Met-tRNAf complex. J. A. Traugh (University of California) described phosphorylase activities isolated from a haemin-stimulated reticulocyte lysate. The diagram below illustrates a working model of how she thinks the removal and addition of phosphate to the initiation factor might be involved in the control of polypeptide initiation.
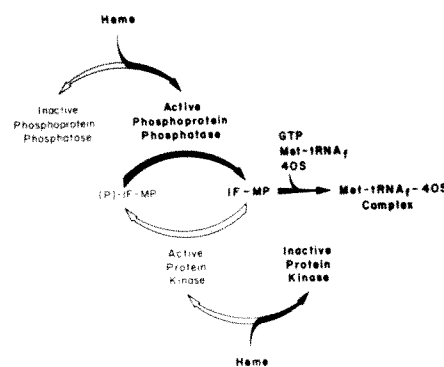
Discussing the action of double-stranded RNA in preventing the translation of encephalomyocarditis virus RNA in an L cell lysate containing interferon, I. M. Kerr (NIMR, London) concluded that there are two steps in the inhibition of translation and hinted that a protein kinase(s) may have a regulatory role in the activation of the inhibitor. In his introduction to the section on interferon and protein synthesis, P. Lengyel (Yale University) pointed out that the effects of interferon were probably many and various. This was borne out by the speakers in this section: no consistent picture of its action emerged.

Developing systems
Some of the best known examples of cytoplasmic control of protein synthesis have been observed in developing systems. mRNA is stored in the cytoplasm and not translated until a further stage in development. Most of the speakers have approached this problem by studying the mRNA before and during its translation to see if there were detectable differences in its properties. C. H. Darnbrough (University of Edinburgh) concluded that, of two populations of messengers with large and small poly(A) size, the latter increased and the former decreased during oogenesis in *Xenopus*. Comparison of the 'active' mRNA isolated from polysomes and the stored RNA from ribonucleoprotein particles, led M. E. Buckingham (Pasteur Institute) to propose that in dividing muscle cells poly(A) tails are involved in maintaining mRNA conformation, rather than

RESULTS obtained by Melton and Giardini (page 309 of this issue of *Nature*) suggesting that oxygen not nitrogen might be the dominant impurity in diamond pose a problem in the light of previous work. In 1936, Robertson *et al.* (*Proc. R. Soc. Lond.,* **A157,** 579–593) classified diamonds into two types by their optical absorption spectra. Kaiser and Bond (*Phys. Rev.,* **115,** 857–863; 1959) using mass spectroscopy, showed that nitrogen impurity was the microscopic basis of this classification; type I diamonds contained nitrogen, type II did not. The nitrogen concentration correlated with the optical absorption lines noted by Robertson *et. al.* Lightowlers and Dean (*Diamond Res.,* 21–25; 1964), using activation analysis, confirmed Kaiser and Bond's results. In some diamonds the nitrogen is paramagnetic, and can be detected unambiguously by spin resonance. Thus there is no doubt that nitrogen is a major impurity in diamond—up to 0.3%—a massive concentration for an impurity in a crystal.

Kaiser and Bond also found small amounts of hydrogen and oxygen. Oxygen has been confirmed by Sellschop (*Diamond Res.,* 35–41; 1975) using activation analysis.

One of the problems in determining the impurity content of diamond is that there are often trapped inclusions of other minerals. These are important, particularly to the geologist, because they indicate the growth environment of the diamond; but they may give a misleading picture of which elements can genuinely enter the diamond lattice as substitutional or interstitial impurities. Kaiser and Bond, and Lightowlers and Dean used stones without visible inclusions. Even so, Sellschop has shown that all diamonds contain submicroscopic inclusions, and these

account for most of the oxygen impurity.

Using mass spectrometry, Melton and Giardini (*Amer. Mineral.,* **59,** 775–782; 1974 and **60,** 413–417; 1975) found that some macroscopic inclusions contain gases—hydrogen, methane, nitrogen, carbon monoxide and water.

Impurities in diamonds

from John Walker

The most abundant element was hydrogen, followed by oxygen and nitrogen. The results are not in conflict with previous studies, because hydrogen cannot be detected by activation analysis, and nitrogen and oxygen need special techniques. The most comparable work is that of Kaiser and Bond, but even here there is no conflict. The latter selected "inclusion-free" diamonds, whereas Melton and Giardini chose crystals containing inclusions, and studied the inclusions.

Thus the picture of impurities in diamond was complicated but consistent. The geologist studies the inclusions; the physicist, the genuine diamond lattice. In particular the physicist, whether he is concerned with optical, electrical, mechanical or thermal properties, knows that type I and type II diamonds have quite different properties, because of their differing nitrogen content.

This consistent picture has been upset by Melton and Giardini's latest paper. They used "inclusion-free" diamonds, and instead of crushing them as previously, they graphitised them at 2,000 °C like Kaiser and Bond. The results are surprising. "The gases released from the diamond, in decreasing

order of abundance, were CO, H_2, H_2O, CO_2, N_2, CH_4, and Ar. The atomic weight percent is O=59.2%, C=28.1%, H=10%, N=2.5% and Ar=0.2%. These data are consistent with the gases released by the crushing of numerous other diamonds *in vacuo*."

Since nitrogen is currently thought to be the dominant impurity in diamond, how can we explain this discrepancy? Melton and Giardini suggest nitrogen contamination from the carbon crucible used by Kaiser and Bond. But this is unlikely because Kaiser and Bond's results are internally consistent—the nitrogen concentration correlated with optical absorption in their crystals. And Lightowlers and Dean, using a different technique, got the same result.

A possible explanation is that the diamonds came from different sources. This is known to affect impurity content. In any case, the fact that Melton and Giardini's diamonds are oxygenrich, surprising though it may be, does not contradict Kaiser and Bond's results on nitrogen. Unfortunately, Melton and Giardini do not give the concentrations of oxygen and nitrogen they found. My estimates based on their data suggest that both lie within the range of accepted values; but it would be better to have the authors' own figures.

One hopes that Melton and Giardini will extend their graphitisation work to diamonds from other sources, and measure the optical spectra before graphitisation, to check Kaiser and Bond's correlations. If the same diamonds could also be studied beforehand by the non-destructive technique of activation analysis, to compare results, the puzzle would probably be resolved. Diamond researchers will await further results with interest.

conferring differential stability. T. Humphreys (University of Hawaii) reported that on fertilisation of sea urchin eggs messenger translation is increased—the messengers are stable but their poly(A) turns over rapidly. Using the same system, G. Giudice (Institute of Comparative Anatomy, Palermo) has shown that although capping does not cause the increase in translational efficiency it may be involved in mRNA selection.

R. A. Laskey (MRC Laboratory of Molecular Biology, Cambridge) has examined the effect on endogenous protein synthesis of injecting polysomes and mRNA into *X. laevis* oocytes. He maintains that oocytes do not have 'extra translational capacity' and that by choosing conditions that accurately reflect the amount of protein synthesised the endogenous protein synthesis decreases competitively with added exogenous mRNA. There is no

competition when the messenger is injected with its own ribosomes. Laskey concludes that the amount of protein synthesised is regulated by a component of normal polysomes and not by messenger availability.

Virus infection
The series of lectures on the control of translation during viral infection was similar to those on masked messengers in development in that many interesting systems were described without providing much insight into the mechanism of control. In general, host mRNA translation is suppressed at the expense of viral protein production. The herpes simplex system was described by C. M. Preston (MRC Virology Unit, Glasgow). He has shown that the effect of this virus on tissue culture cells can be reproduced in a lysate of the cells. The lysate can also respond to exogenous mRNA, but its response is decreased

if the lysate is prepared from infected cells. The lesion can be traced to the ribosomes of the infected cells, since proteins washed off reticulocyte ribosomes by salt treatment restore the translational capacity. R. E. Thach (Washington University) has shown that EMC RNA can 'out-compete' host mRNAs but that the competition is relieved by an excess of the initiation factor corresponding to rabbit IF-E6 (IF-M3). L. Carrasco (ICRF, London) has found that after infection with picornavirus, changes take place in the cell membrane which impair the sodium transport system, leading to an increase in the cell concentration of sodium ions and a decrease in potassium. Experimental evidence has confirmed his suggestion that increased sodium concentration favours the translation of viral, rather than host, mRNAs. It may be that at these altered salt conditions binding of initia-

tion factors to viral messengers is favoured—hence their preferential translation.

In the section dealing with the topography of protein synthesis G. Blobel (Rockefeller University, New York) presented a very appealing model to explain the translation of mRNA for secretory proteins on membrane-bound ribosomes. It requires that mRNAs for secretory proteins code for an N-terminal peptide that is recognised by the membrane, resulting in the attachment of ribosomes to the membrane so that the growing protein is vectorially transported into the intracisternal space. This is illustrated in the diagram : the numbers indicate sequential steps in the synthesis of the protein. Bernard Mechler (University of Cambridge) proposes a slightly different scheme since he has to explain his observation that the binding of ribosomes to the membrane can be mediated by the mRNA directly

rather than its translational product. H. Lodish (MIT, Cambridge) showed that translation of the polycistronic Sindbis virus RNA can take place in the cytoplasm or on membranes using the same ribosomes but producing different proteins. He concludes that the N-terminus of the nascent chain
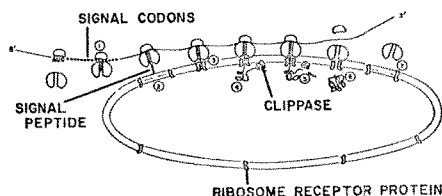
must specify the binding of ribosomes to membranes.

The participants left the conference with the impression of many examples of cytoplasmic control of protein synthesis but few explanations of the mechanisms involved, an admirable situation sure to keep them in business for several years to come. □

Flow of superfluid ³He

from P. V. E. McClintock

To push some of the newly discovered superfluid ³He through a fine tube, in order to see how it behaved, might seem a particularly obvious experiment, were it not for the daunting difficulties inherent in working at temperatures near 2 mK. In fact, it is a quite remarkable achievement in terms of experimental design and technique that has enabled R. M. Mueller, E. B. Flint and E. D. Adams to report (*Phys. Rev. Lett.*, **36**, 1460; 1976) the first studies of equilibrium flow phenomena, only four years after observation of the superfluid phases was first suspected. Their experiments were carried out in the Physics Department at the University of Florida.

Because of extreme difficulties of making thermal contact in this ultralow temperature range, it is usually necessary to design the cryogenics and the experimental cell as an integral unit. Accordingly, the Florida group have devised an ingenious arrangement whereby a pair of Pomeranchuk cells was used both for cooling the ³He into the superfluid A-phase and also for inducing a flow of liquid through the narrow tube which connected them together.

The Pomeranchuk cooling technique depends on the fact that, at very low temperatures, ³He gets colder if it is solidified by compression. Thus, by applying a mechanical force to reduce the volume of a sample of ³He enclosed in a flexible chamber, thereby solidifying some of it, the remaining liquid may be cooled through the superfluid

transition at 2.6 mK. The mechanical force is usually generated by way of a separate hydraulic system in which the working fluid is liquid ⁴He.

The two Pomeranchuk cells of the Florida apparatus were fitted with completely separate hydraulic systems, so that their volumes could be varied independently of each other. By compressing both cells simultaneously the ³He could be cooled without inducing any movement of liquid in the connecting tube. On the other hand, if one cell was compressed while the other was being expanded, then liquid could be made to flow through the tube at constant temperature. A nuclear magnetic resonance (NMR) technique was used for monitoring events in the flow tube.

In view of a calculation by Fetter (*Phys. Lett.*, **54A**, 63; 1975), it had been anticipated that the flow would result in a shift downwards of the NMR resonant frequency. It was surprising, therefore, that rather than changing its frequency, the resonance line in practice developed satellites, and then abruptly disappeared, as the flow velocity was increased. Although the behaviour was not always reproducible in detail, it was found that the satellite structure never appeared when the external magnetic field was applied at 90° to the flow direction. This may perhaps be related to Fetter's prediction that no frequency shift should be seen in this configuration. A further complication was that, before its final rapid fall, the resonance signal sometimes

diminished in a series of steps as the flow velocity was increased.

The most definite and interesting observation, however, was the remarkably small velocity at which the signal disappeared. The exact value depended on temperature, but fell within the range 0.003–0.02 mm s⁻¹. These velocities are several orders of magnitude smaller than the Landau critical velocity at which the superfluidity would be expected to disappear as a result of pair-breaking, so that some other mechanism must presumably be involved. It is interesting that investigations of the heat conduction by R. T. Johnson, R. L. Kleinberg, R. A. Webb and J. C. Wheatley (*J. Low Temp. Phys.*, **18**, 501; 1975) also yielded multiple critical velocities for the superfluid which were remarkably small (0.2–0.8 mm s⁻¹), although considerably larger than the velocities reported here. The authors demonstrate, however, that the two sets of measurements may in fact be in quantitative agreement if it is assumed that the flow did not occur through a transfer of bulk liquid: if only the superfluid component of the liquid had moved, leaving the normal fluid component clamped in the tube by its own viscosity, it would have had to move relatively briskly to simulate a given rate of apparent bulk flow because it constitutes only a small proportion of the liquid at the temperatures in question. On this basis they deduce superflow critical velocities in the range 0.14–0.64 mm s⁻¹, in good agreement with those derived from the thermal conduction experiments.

Why should these critical velocities be so small? One is immediately reminded of the parallel situation in superfluid ⁴He, where the actual critical velocities for superflow are usually about 1 mm s⁻¹, compared with the much larger Landau critical velocity for roton creation of around 50 m s⁻¹. In the case of ⁴He, the discrepancy has been shown to arise from the relative ease with which quantised vortex lines may be generated in the liquid for flow velocities far below the Landau velocity. It seems very likely that something of a similar nature is happening in ³He. The situation is, of course, infinitely more complicated because the liquid is governed by a vector order parameter, is anisotropic, and takes up complicated textures depending on the shape of the container, applied electric and magnetic fields, and so on. Furthermore, it is not yet by any means clear what would be the analogue of a vortex, in the case of ³He: there may in fact be several different sorts of collective excitation of this general nature.

Assuming, however, that the low critical velocities observed in these experiments arise from the onset of

dissipative processes, it seems probable that at least one completely new type of collective excitation is being created, thus promising to add yet another whole dimension of complexity to our understanding of this extraordinarily complicated liquid. □

Two approaches to gene synthesis

from Maria Szekely

THE recent discovery that synthetic genes can be used to produce genetic recombinants (see *News and Views,* **260,** 189; 1976) opens wide possibilities in the field of genetic engineering. Well-characterised DNA sequences can be inserted into plasmid DNA thus providing specific probes with known sequences for studying the organisation and expression of eukaryotic genomes. Amplification of the inserted gene yields large amounts of specific sequences which can be used for the production or isolation of highly labelled mRNAs, for sequence studies, and so on. At the same time, the safety risk is lessened, as no unknown, harmful genetic information can be integrated into the recombinant.

So far only one eukaryotic gene, that of β-globin, has been inserted into bacterial plasmids. Almost simultaneously, four laboratories reported the production of recombinants from β-globin DNA, using different techniques but all of them starting with reverse transcription of globin mRNA into cDNA. (Rougeon *et al., Nucleic Acids Res.,* **2,** 2365, 1975; Efstratiadis *et al., Cell,* **7,** 279, 1976; Rabbits, *Nature,* **260,** 221, 1976; Higuchi *et al., Abstract ICN-UCLA Winter Conference,* 1976). The advantages of this technique are obvious, there are, however, also some limitations to the method. It follows from the principle of the technique that the DNA inserted into the plasmid can contain only the sequences present in the mRNA : any additional, non-transcribed sequences that may be present in the original gene will be lost. It remains to be seen how their absence affects the function of the integrated gene. In practice, the limitations go even a step further : it depends on the efficiency of the reverse transcriptase how much of the genetic information present in the mRNA will be passed on to the synthetic gene. In many laboratories only fragments of cDNA copies could be obtained with this enzyme. Rougeon *et al.* used a 300–400 nucleotide long cDNA preparation to synthesise the double-stranded globin DNA and after amplification ended up with a 100 to 150

base-pair long integrated globin DNA, about one fifth of the length of the mRNA.

Maniatis' group reported last year the synthesis of a full-length copy of globin mRNA (Efstratiadis *et al., Cell,* **4,** 367; 1975). A few months ago the same group described the sythesis of double-stranded globin DNA with the aid of DNA polymerase I, making use of the hairpin loop structure at the 3′-end of this cDNA, which served as primer in the synthesis of the anti-cDNA strand. (Efstratiadis *et al., Cell,* **7,** 279; 1976). They succeeded in inserting this synthetic gene into PMB 9 plasmid DNA and in their recent paper (Maniatis, Kee, Efstratiadis and Kafatos, *Cell,* **8,** 163; 1976) they describe in detail the techniques of integration of the DNA, transformation of *Escherichia coli* cells and identification of the recombinant.

The synthetic gene is 580 base pairs long, about 80 nucleotides shorter than β-globin mRNA. It contains the full coding region plus 40 and 110 nucleotides of the untranslated regions at the 5′ and 3′ termini, respectively. Molecular hybrids were constructed by annealing the globin DNA which carried poly(dT) tails to *Eco*RI-treated plasmid DNA to which poly(dA) tails have been attached. Transformation of *E. coli* HB101 with this hybrid DNA yielded 600 tetracycline-resistant colonies, a high proportion of which contained sequences of the β-globin gene, as was shown in hybridisation assays. The size of the inserted DNA did not change upon construction of the molecular hybrid, transformation and cloning.

The method seems suitable for general application in the synthesis and amplification of eukaryotic genes. Maniatis' group has already synthesised a number of other genes by copying eukaryotic mRNAs.

It is essential for the usefulness of such techniques that the original structure of the gene should be retained in the recombinant. Fidelity in copying the mRNA and stability of the DNA structure during integration and transformation is therefore of vital importance. Maniatis *et al.* characterised the synthetic gene and the inserted globin DNA after amplification by comparing restriction maps of the DNAs and correlating them to the structure of globin mRNA. Using eight restriction enzymes with known recognition sequences, the restriction map of synthetic DNA was established and compared with the cleavage pattern predicted from the nucleotide sequence of β-globin mRNA. As only part of the nucleotide sequence has been determined so far, some cleavage sites were established by deducing the nucleotide sequence from the known amino acid sequence of β-

globin. The restriction map of synthetic globin DNA was in perfect agreement with that expected from the mRNA structure.

The same restriction enzymes were used individually and in combination on the recombinant DNA. By using *Hha* endonuclease, an enzyme which does not cleave synthetic globin DNA but produces many fragments from PMB 9 DNA, one fragment, carrying the inserted globin gene could be isolated and characterised. The restriction map of this fragment was established, and the orientation of the inserted DNA and the distance between the restriction sites determined. These proved to be the same as in the synthetic gene before integration, confirming that the entire synthetic globin DNA molecule was amplified without sequence rearrangements. Some slight changes were detected in the poly(dA)-poly(dT) tails only.

The approach used by Maniatis' group yields very useful DNA probes which will greatly enhance progress in our understanding of the structure and function of the eukaryotic genome. The success of a different, more chemical approach has been reported recently from Khorana's laboratory : the complete gene of suppressor tyrosine tRNA of *E. coli* has been synthesised without using an RNA template. The techniques were worked out years ago by Khorana's group. Following a known nucleotide sequence, chemical synthesis is used to produce blocks of nucleotides and the ligase reaction to join them. The synthetic gene contains the total sequence of the precursor tRNA and includes the sequence of the promoter site and of the stop signal for transcription. Twenty-four scientists contributed to its synthesis which has taken 9 years. This gene has now been introduced into an *E. coli* amber mutant by way of a transducing phage (the details of the technique have not been published so far) and it was found to function correctly as a suppressor gene in the bacterial cell.

Khorana's method requires more time and work than the copying of an RNA template and its use is limited to the synthesis of genes the sequence of which has been fully established. Still, this approach has some great advantages over the reverse transcription of RNAs. The gene synthesised by Khorana contains also sequences which are not present in the final gene product, sequences lost during processing of the RNA and non-transcribed control signals. Such sequences cannot be incorporated into a reverse transcript. The purely synthetic method also makes it possible to introduce specific alterations into the gene structure and to study the effect of such "mutations". It would be interesting to

see if the two approaches could be combined and a new technique developed which would unite the relative simplicity of the copying technique with the further possibilities provided by the methods of chemical synthesis. □

New method for measuring nuclear lifetimes

from P. E. Hodgson

AN ingenious method for measuring very short nuclear lifetimes has recently been developed that greatly extends the range covered by existing techniques. This is an important advance because these lifetimes can often be calculated from nuclear wavefunctions, and thus provide a sensitive test of the theories used to calculate the wavefunctions.

Nuclear lifetimes are usually measured by the delayed coincidence and the Doppler shift techniques, and these cover a wide range of times greater than about 10^{-15} s. For shorter lifetimes there has hitherto been only the blocking technique which is restricted to a few suitable materials. The new method seems likely to provide a way of measuring lifetimes from about 2×10^{-15} to 6×10^{-18} s, and thus extends the region of measurable lifetimes by more than two orders of magnitude.

The new method is based on a very simple idea. If a nucleus of charge Z decays by electron capture to an excited state of another nucleus of charge $(Z-1)$ a vacancy is left in the atomic K-shell from which the electron has been captured. This vacancy will soon be filled by a transition from a higher atomic state with the emission

Histogram showing the spectrum of all X rays in coincidence with all decay protons. The curves show the spectra measured independently for the As and Ge decays.

of an X ray. Now if the X-ray transition takes place before the nuclear state decays, its frequency is characteristic of the nucleus of charge $(Z-1)$. But if the nucleus decays by proton emission before the X-ray transition takes place, then the X ray will have a frequency characteristic of the nucleus of charge $(Z-2)$. Since these X-ray frequencies can easily be measured, it is easy to find out whether the nuclear decay took place before or after the atomic decay. In practice there is of course a distribution of decay times for both processes, so it is the relative proportion of the two types of decays that is measured and this gives the ratio of the nuclear and atomic lifetimes. Since the atomic lifetimes are well known both experimentally and theoretically, the nuclear lifetime is found.

This method has been devised by a group of scientists from the Chalk River Laboratories and the University of Toronto in Canada. They have obtained data on the lifetimes of ⁶⁵Ga, ⁶⁹As, ⁷³Br, and ⁷⁷Rb, and in their first publication (*Phys. Rev. Lett.*, **37**, 133; 1976) they give details of their result for ⁶⁹As.

In this case the decay sequence starts with $^{69}_{34}$Se undergoing electron capture to an excited state of $^{69}_{33}$As, which in turn undergoes proton emission to a state of $^{68}_{32}$Ge.

The initial nucleus ⁶⁹Se was produced in the ⁴⁰Ca(³²S,2pn)⁶⁹Se reaction using the 100 MeV beam of the Chalk River tandem Van de Graaff accelerator. After irradiation, the target was moved rapidly to a counting position where arrays of counters measured the X rays from the atomic transitions, the delayed protons from the decay of the ⁶⁹As to ⁶⁸Ge, and the gamma rays from the nuclear transitions in As.

The histogram in the figure shows the energy spectrum of all the X rays in coincidence with the delayed protons, compared with normalised curves corresponding to the decays of Ge and As found in independent measurements. It is clear that the histogram can be very well fitted by the sum of these two curves.

Many states of ⁶⁹As can be reached by electron capture, so the decay protons are themselves varying in energy. Since the relation between the relative numbers of Ge and As X rays depends on the proton energy, so do the lifetimes of the states of As. The lifetimes of the states in As vary from 2×10^{-16} to 10×10^{-16} s. These numbers are average lifetimes for successive bands of states and rather detailed calculations based on the statistical theory of nuclei are required to interpret them in detail.

If there had been only one state in As fed by the electron capture in Se the experiment would have given an

unambiguous determination of its lifetime. The technique is clearly successful and is likely to be applied to determine the lifetimes of many states, and hence to increase our knowledge of nuclear structure. □

Effects of a long hot summer

from Peter D. Moore

THE unusually warm and dry summer experienced in Britain this year has undoubtedly influenced our vegetative cover. Apart from widespread damage by fire, the high temperatures and low water availability have themselves produced some very evident changes. Roadside verges and lawns are parched, with only the deep-rooted rosette species showing signs of life. Among trees, birch appears particularly susceptible and dead birches are now abundant on heathlands and in mixed woodland both in the lowland and the upland zones of the country.

Many of the common species affected will probably recover quickly and even if they do not, any change in status will be difficult, if not impossible, to assess. There are certain species, however, which may merit especial attention at this time, such as those which lie at their climatic limit in Britain and which may expand or contract their range as a result of the effects of this and preceeding summers. There are also those species which are restricted to lower altitudes in the British Isles because of the shorter growing season in higher regions; they may attain new heights. And there are those arctic alpine species which are limited by high summer temperatures and which may, therefore, be suffering.

One species which reaches its northwestern limit of distribution in the British Isles and which might be expected to respond to the present conditions is the stemless thistle (*Cirsium acaulon*). Pigott (in *The Flora of a Changing Britain*, edit. by Perring, F., 32, Classey, Hampton, Middlesex, 1970) has shown that the temperature of the flowering head of this plant is critical in determining the rate of development of the embryo. Only in southern and eastern Britain does the plant normally produce viable seed, though some outlying populations in the north and west appear to have been established during exceptionally warm summers.

Many plants are common at fairly low altitudes and yet fail to flourish on mountain summits because of their inability to grow adequately, or flower, or produce mature, viable seed, within
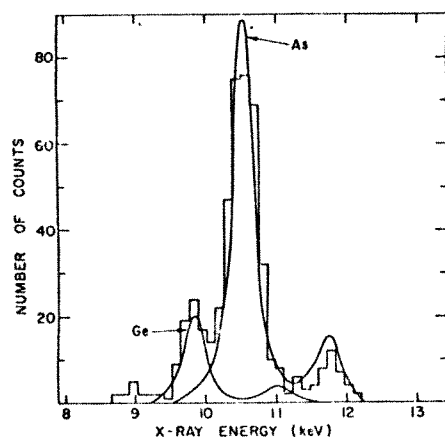
NEW observations of H-β or Lyman-α fluxes of a sample of 57 Seyfert galaxies and QSOs suggest strongly that QSOs are indeed distant, bright Seyfert galaxies too remote for the faint outer structure of the galaxies to be observed. As well as confirming the now increasingly accepted view that the redshifts of QSOs are cosmological distance indicators, the suggestion that QSOs are high-redshift Seyferts provides a new use of these objects as cosmological probes, and suggests some limits on the so-called "deceleration parameter" of our Universe.

The possibility of a link between active galaxies such as Seyferts and QSOs has long been recognised in a qualitative sense, two touted possibilities being that explosions on the scale of a QSO may occur in the nuclei of all spiral galaxies from time to time, or that such violent activity may mark a step in the evolution of a proto-galaxy, after which it settles down into a quiet spiral with no further comparable activity. Supporters of such ideas have had powerful circumstantial evidence to hand in the observation that class I Seyferts and QSOs cannot be distinguished from their spectra alone. Both have broad Balmer emission lines, narrow forbidden lines and power-law spectra, the difference being that Seyferts are seen as active nuclei surrounded by visible galactic material, while QSOs have higher redshifts and no surrounding galaxies are seen. Because of the confusion effects from bright nuclei, surrounding galactic disks could not, however, be observed at redshift greater than 0.1—so any Seyfert of class I with $z>0.1$ would be identified automatically as a QSO.

Weedman has now developed from this observation a quantitative

Seyfert galaxies and QSOs

from John Gribbin

measure of the similarity between the two kinds of object (*Astrophys. J.*, **208**, 30–36; 1976). Earlier studies along these lines had compared the fluxes from the two in a specified part of the observed spectrum; as Weedman points out, this is not the best approach because the redshift effect means that different parts of the emitted spectra are shifted into the chosen observational window. Instead, he has made measurements of the emission of his sample of Seyferts and QSOs in the hydrogen lines, H-β or Lyman-α. The H-β lines were measured wherever possible, but at large redshift these are shifted out of the visible range and for those objects L-α can be measured instead, the relative strength of the two lines (based on model calculations by K. Davidson, *Astrophys. J.*, **171**, 213; 1972) then giving an estimate of the strength of the invisible H-β.

The results show clearly a continuous smooth distribution of luminosity in line with the cosmological interpretation of the redshifts of these objects. The range covered, equivalent to 11.5 magnitudes, is large, but as Weedman points out still less than the luminosity range of main sequence stars. Accepting this as confirmation that QSOs are simply bright class I Seyferts, this immediately provides a new probe of the Universe extending over a wide range of redshifts, and Weedman points out some preliminary conclusions from this which set limits on some cosmological parameters. In particular, the luminosities of high redshift QSOs (which are now regarded as high-redshift Seyferts) are relatively faint, suggesting that either there is some limit to the brightness of these objects or that the deceleration parameter of the Universe, q_0, is negative. This tentative first conclusion is, however, less important than the fact that at last it seems possible that QSOs can be used with confidence as genuine cosmological probes to define the structure of our Universe on the large scale, and to rule out some of the large family of possible models now applied to the Universe. The problem of the energy source of QSOs remains, of course (even if there is some upper limit), but here too it seems that this problem can now be approached from a new and promising direction by investigating the more easily studied and nearer members of the family, the Seyfert galaxies.

---

a short growing season. Pearsall (*Mountains and Moorlands*, 49, Collins, 1956) described his observations on the heath rush (*Juncus squarrosus*) and showed how the length of the flowering stalk, the number of flowers per inflorescence and the number of viable seeds produced per capsule are all inversely related to altitude. The production of viable seed for this species is uncommon at altitudes over 820 m. It is also reported, however, that the heath rush produced seed on the summit of Ben Wyvis (1,036 m) in the unusually warm summer of 1947.

Among lowland plants, some occur only in those parts of the country experiencing high summer temperatures. For example, the round-headed rampion (*Phyteuma tenerum*) occurs mainly within the 16.6 °C July average isotherm. Such species may also be expected to extend their range under current conditions.

Other species may find high summer temperatures detrimental, particularly species of arctic alpine distribution. Dahl (*Oikos*, 3, 22; 1951) observed that

arctic alpine species grown in lowland gardens suffered most during hot summers. He has also demonstrated that many of these species have distributional patterns in which their limits coincide with low summer temperatures. Similar observations have been made recently for certain arctic alpines in Newfoundland (Damman, *Can. J. Bot.*, 54, 1561; 1976), for example, the bog bilberry (*Vaccinium uliginosum*) is found only in areas having an average temperature in the growing season below 15 °C. It is known that dark respiration rates of many arctic alpine species are higher than their lowland counterparts (Billings and Mooney, *Biol. Rev.*, 43, 481; 1968) and since respiration is stimulated by high temperatures, the plant could experience difficulty in maintaining a positive carbon balance during a hot summer.

One species which does appear to be suffering in upland Britain this year is the bilberry (*Vaccinium myrtillus*). This is a species which dominates some open moorlands in the uplands of western Britain, yet is found only under a

woodland canopy in the east and in continental Europe. Microclimatic factors, possibly involving temperature and humidity may therefore be critical for its survival. This summer the open populations of the west are showing evident signs of stress when unprotected by a tree cover. It is possible that its reduced vigour will allow such species as the heather (*Calluna vulgaris*) to spread at its expense.

Our freak summer offers valuable opportunities to increase our understanding of many plant species and one can only hope that any ensuing changes in vegetation will not remain unrecorded. This may be an appropriate time to mobilise that large and willing labour force, the amateur botanists, to undertake a census of selected plant species on the model of the British Trust for Ornithology's common bird census. Perhaps the Botanical Society of the British Isles is the appropriate organisation to initiate such a census which could begin to provide an answer to the question how responsive and fluid is our flora? □

Structure and function of haemocyanin

from J. V. Bannister and H. A. O. Hill

The V Haemocyanin Workshop sponsored by the European Molecular Biology Organisation was held on August 1–4, 1976, at the new Marine Biology Laboratory of the University of Malta. The meeting was organised by the Department of Physiology and Biochemistry of the University.

THE haemocyanins are multisubunit copper proteins with respiratory function found in arthropods and molluscs. The molecules are of giant size with a molecular weight of up to 1 million in arthropods, 4.5 million in cephalopods and 9 million in gastropods. The stoichiometry of oxygen binding is one O_2 per two Cu and the copper content indicates a minimum functional subunit of 75,000 in arthropod and 50,000 in molluscan haemocyanin. Subunits of 75,000 molecular weight appear to exist in arthropod haemocyanin and can be regarded as monomers. The monomer molecular weight in molluscan haemocyanin has not been fully resolved.

The study of the structure and behaviour of these extra-large molecules is a matter of real importance, providing as it does a stepping stone to the understanding of the regulation and assembly of still larger biological structures, such as microtubules or ribosomes (J. Wyman, University of Rome). Three-dimensional image reconstruction from electron micrographs has provided a model for gastropod haemocyanins (Mellema and Klug, *Nature*, **239**, 146; 1972) which showed the molecule to consist of a hollow cylindrical drum closed at both ends with material of fivefold symmetry ("collar") containing a "cap" in the centre. The wall has approximately tenfold rotational symmetry and consists of 60 quasi-equivalent morphological units bounded by two dominant sets of helical grooves. E. F. J. Van Bruggen (University of Groningen) presented results on the possible folding of a long polypeptide chain in the strongly curved wall segments from a half molecule. A segment would accommodate one-twentieth of the molecule of *Helix pomatia* haemocyanin and one-fifth of the fivefold "collar". The "cap" of Mellema and Klug is thought to be an artefact. It has proved difficult to dissociate gastropod haemocyanins into material smaller than one-twentieth of the molecule (about 450,000 molecular weight) without breakage of peptide bonds. The twentieths have therefore been regarded as the monomers of these haemocyanins. If so a monomer contains about nine oxygen binding sites or "domains". C. Gielens (University of Leuven) showed that two sites between the domains are susceptible to proteolysis. Various proteolytic enzymes including subtilisin, papain and Pronase give 50,000 molecular weight fragments with whelk, *Murex trunculus* haemocyanin (J. Bannister, University of Malta). Similar fragments were observed after prolonged reduction followed by carboxymethylation of the haemocyanin. The protomer may therefore be as small as a single O_2 binding domain in this haemocyanin.

The structure of arthropod haemocyanins seem less perplexing. The assembly of 60S molecules in *Limulus*, and 34S in scorpion and tarantula haemocyanins, from 16S subunits was discussed (E. F. J. Van Bruggen, University of Groningen; J. Lamy, University of Tours; B. Linzen and R. Loewe, University of Munich). Each 16S particle contains six monomers (5S). B. Salvato (University of Padova) questioned this structure of the 16S subunit in *Limulus* haemocyanin. W. Love (John Hopkins University) reported the conditions in which whole molecules of *Limulus* haemocyanin and some of its subunits can be crystallised. The crystal forms from which preliminary X-ray data were obtained are still unsuitable for high resolution structure analysis and conditions are being sought for the growth of more suitable crystals. The 60S whole molecule is predicted by X-ray diffraction to consist of about 50 subunits of which there are at least nine different kinds present in unknown ratio.

Limulus haemocyanin can be dissociated into five distinct subunits (I–V) of about 66,000 molecular weight. The fully associated molecule is therefore a 48-meric structure. Van Bruggen discussed the association of subunits I–V as seen in electron micrographs. The 5S structure is a roughly spherical protein particle with a diameter of about 65 Å. None of the separated subunit fractions was found to be capable of forming an intact 60S structure by itself. The 60S structure is formed by the stacking of two-bridged tetramers probably in a staggered arrangement. The heterogeneity of the 5S subunits was demonstrated in scorpion haemocyanin by Lamy. Six subunits could be isolated from the haemocyanin, one of which—subunit 4—was shown to have a particular ability to polymerise in one, probably hexameric 16S particle.

Until now the O_2 bound in haemocyanin has been considered to form a peroxo complex with two Cu(II) atoms. R. Lontie (Leuven) proposed consideration of a Cu(I) . . . Cu(III) complex. X-ray photoelectron spectroscopy, however, indicates that Cu in oxy-haemocyanin is cupric because the $2p_{3/2}$ and $3p_{3/2}$ binding energies of the Cu are similar to those in Cu–Zn superoxide dismutase (H. Van der Deen, Groningen).

Oxyhaemocyanins have a strong absorption band in the near ultraviolet. Photo-oxidation experiments with light at the absorption wavelength show that tryptophan is located near the copper-oxygen site (G. Jori, Padova). M. De Ley (Leuven) proposed that photo-oxidation of haemocyanin results in the activation of oxygen by tryptophan.

Haemocyanins show both a reverse and a normal Bohr effect. Cooperativity is observed with whole or half molecules and is mediated by divalent ions or high ionic strength. The O_2 equilibrium and kinetics of gastropod and arthropod haemocyanins were discussed. Undissociated *Panulirus interruptus* haemocyanin under conditions of co-operative O_2 binding was reported to show a non-autocatalytic oxygen dissociation behaviour (H. Kuiper, Groningen). Temperature-jump studies showed that the oxygenation reaction is heterogeneous in *Limulus* haemocyanin and its isolated subunits (C. Bonaventura, Duke University Marine Laboratory). The reverse Bohr effect in *Buccinum* haemocyanin was investigated by E. Wood (University of Leeds). The kinetic explanation for this effect was reported to be that the dissociation velocity constant was strongly *p*H-dependent whereas the association explanation was reported for a positive Bohr effect except that the *p*H dependency of the velocity constants goes in the reverse direction (M. Brunori, University of Rome).

The question of whether molluscan and arthropod haemocyanins are similar because of convergence or are homologous and have descended from a common ancestral gene remains unanswered. A. Ghiretti-Magaldi (Padova) showed that the available amino acid compositions of several arthropod and molluscan haemocyanins can be used to reconstruct a consistent evolutionary tree.

Since the first Haemocyanin Workshop in Naples 10 years ago, much progress has been achieved but nevertheless, the protein is proving to be more difficult to study than was first imagined. ☐

articles

Cyclic climatic variations in climate over the past 5,500 yr reflected in raised bogs

Bent Aaby

Geological Survey of Denmark, Thoravej 31, DK-2400 Copenhagen NV, Denmark

Investigations of Danish raised bogs apparently indicate cyclic long term climatic variations with a periodicity of about 260 yr over the past 5,500 yr. The result could be used for modelling future climatic trends.

RAISED bogs represent a special type of peat bog. they are lens shaped, and highest in the centre. They receive no minerogenic water from their surroundings, but have their own water regime, with all of the moisture supplied from the atmosphere, as rain or snow. The humidity in the upper peat layers depends, therefore, essentially on precipitation, temperature and evaporation, a fact which makes raised bogs very sensitive to variations in climate.

Though past climatic changes are reflected in raised bogs as variations in the degree of decomposition (or humification) of the peat, interpretations are not straightforward. For example, a trend towards accumulation of more light coloured peat (decreasing humification) indicates increasing humidity in the bog, which may be attributable either to higher precipitation or lower temperature, or to a combination of the two. It is not possible to identify which of the two parameters is responsible, or to tell exactly how drastic or persistant a recorded climatic change may have been. The opposite shift, from light to dark coloured peat, does not necessarily depend only on climatic parameters, because peat formed under stable climatic conditions will also show the same trend in decompositional behaviour, since the increase in the relative distance to the water table changes as new peat accumulates. Shifts to drier bog conditions are therefore ignored as indicators of climatic changes.

Although changes from dark to light peat formation may indicate climatic shifts, local conditions, such as surface drainage patterns, may also affect patterns of peat formation. It is therefore important to investigate large open peat sections to distinguish widespread alterations (to be found at three or more places in a section) from local alterations. The surface of raised bogs from suboceanic areas comprises drier hummocks and wetter hollows. General variations in the water regime often cause an alteration in their distribution pattern (Fig. 1).

From the rather stationary centre the hummocks transgress in stable and/or dry periods and regress when conditions become generally wetter. A theoretical section, based on the Draved bog profile is shown in Fig. 1B. It can be demonstrated that Draved hummocks have existed at the same place for more than 2,500 yr, contrary to the theory of cyclical hollows/hummock succession (see, for example, ref. 1). Accordingly, an autonome succession between hollows and hummocks cannot have been responsible for the observed shifts in peat formation.

The hollow peat is more likely to reveal a detailed record

than the hummocks, where smaller variations in the humidity are not registered. Levels corresponding to the transgression and regression stages of the hummocks can be detected by measuring humification variations in the hollows.

Obtaining the material

I investigated a number of large, open peat sections in five raised bogs in Denmark. All of the bogs cover a rather extended area (more than 200 ha), and the sections are situated in the central part of the bogs. The degree of humification was measured by the colorimetric determination (see ref. 2). A humification curve is shown in Fig. 2, together with the relative frequency of two rhizopod genera. These genera occur in raised bogs connected to the moist part of the bog surface. Population frequencies reinforce indications of past changes in the water regime on the bog, as reflected in the humification curve.

Levels reflecting past climatic changes have been dated using the radiocarbon method. From Draved Mose a time/depth scale[2] based on 55 radiocarbon dates has been used. In the other cases one or two samples above and beneath levels in question have been dated. The ^{14}C dates have further been calibrated to calendar years in accordance with the American bristlecone pine chronology (see ref. 3).

Figure 3 shows the dates of climatic shifts as determined in Danish bogs. Some dates are identical within statistical uncertainty, reflecting the general character of the climatic events. For example, the onset of colder and/or more humid weather conditions around AD 500 has been recognised in four different bogs[4]. The dates seem to be systematically distributed in time. Generally, there is a time interval of around 260 yr between palaeoclimatic shifts, as indicated in the left-hand side of the figure; but sometimes there is the double time span, 520 yr, between shifts. The reason could be that the intermediate climatic changes were too weak to be reflected in the bogs (for example, the missing shift around AD 1800, see Fig. 4), or that data are too sparse to show missing events. A few dates seem not to accord with the supposed periodicity (at ~ 1400 BC and ~ 2000 BC).

Although the raised bogs do reflect past climate, not every climatic shift is indicated in the peat structure because the rate of decay is rather slow. The degree of decomposition may therefore rather depend on general environmental conditions during an extended period[5]. When light, weakly decomposed peat was formed, wet periods were more frequent than dry ones, while opposite conditions prevailed when dark peat formation predominated. Raised bogs are therefore supposed to react as biological low pass filters only reflecting general long term tendencies
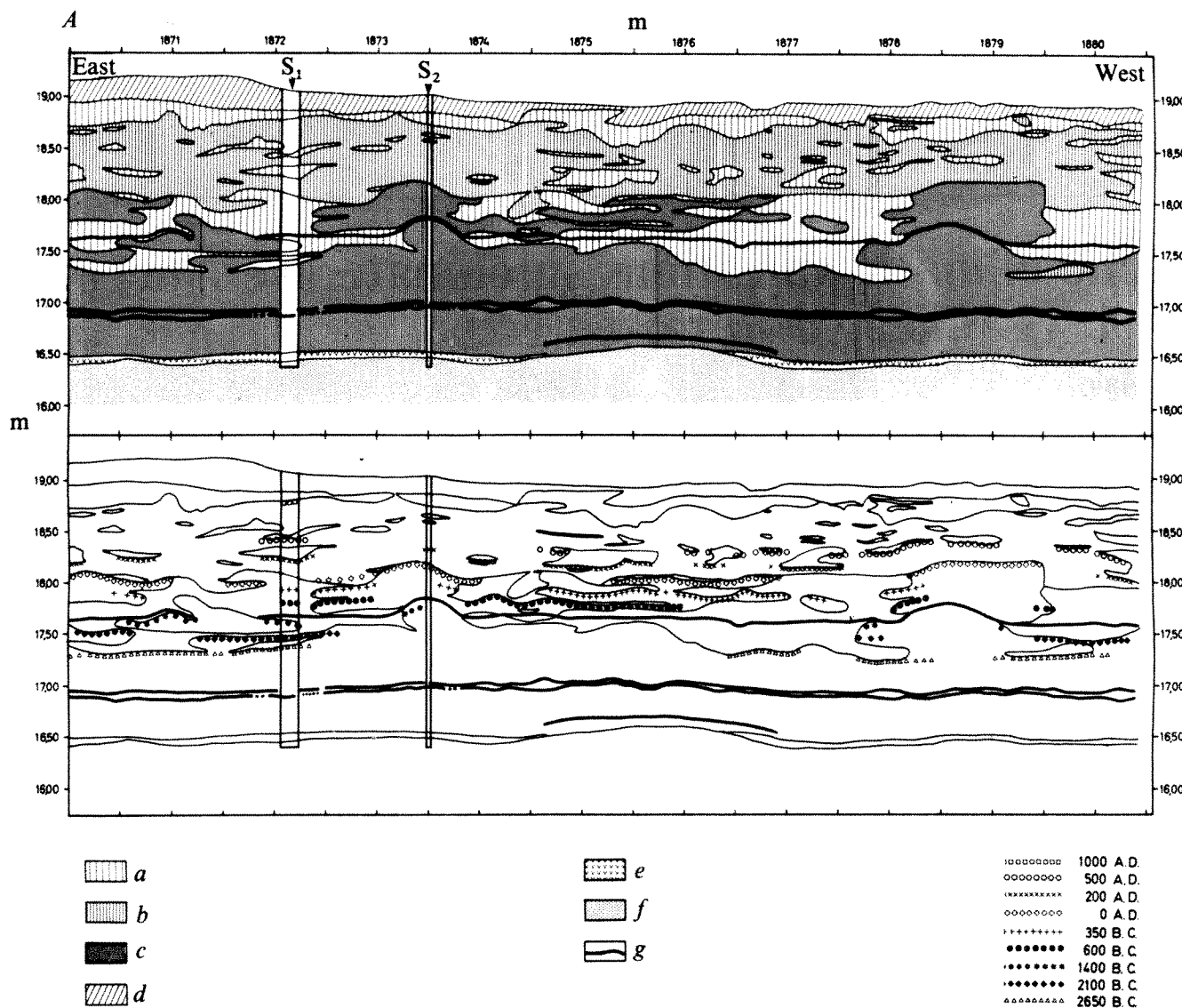
Fig. 1 *A*, Part of a 30-m long peat section from the central part of Draved Mose. The bog lies on aeolian sand. Peat formation began about 5500 BC. *a*, Slightly humified, sphagnum–vaginatum peat; *b*, moderately humified sphagnum–vaginum peat; *c*, highly humified sphagnum–vaginum peat; *d*, disturbed peat layer; *e*, hydromor; *f*, sand; *g*, charcoal layer; *s*, sampling sites. Top, visual variations in peat colour, together with distinct charred layers. Highly humified peat (*c*) dominates most of the section at about 1,873.5 m (at sampling site S₂) and at 1,879.0 m, where two hummocks have existed for more than 2,500 yr. From the rather stable centre the hummocks transgress in dry periods and regress when it becomes generally wetter (see *B*). Bottom, Approximately synchronous levels. Based on the three distinct charred layers (two lying close together at 16.90–16.95 m and one at 17.60 m) and 55 radiocarbon dates from sampling site S₁, the bog development in this area has been reconstructed. General changes from dark to light peat formation occur nine times (see the symbols). The dates of changes are given rounded off to
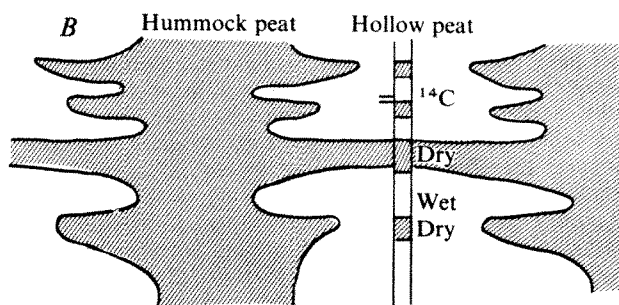
the nearest 50 yr. *B*, Theoretical peat section, showing that the hummock area increases in dry periods, and becomes smaller in wetter periods.

in climate, and the 260-yr cycle seems to be the shortest periodicity revealed.

This interpretation of changes in the degree of humification is supported by historical evidence from the past millennium. From England a temperature curve for the past 1,100 yr is available (Fig. 4). If the curve is smoothed by a 200-yr low pass filter it shows the same general climatic trend reflected in the Danish bogs. The latest period of colder/wetter climate previously detected from raised bogs occurred during the first part of the sixteenth century, that is, at the beginning of a temperature decline corresponding to the onset of the "little ice age". The preceding change in peat formation patterns (in the middle of the thirteenth

century) seems to correspond to the end of the mediaeval warm periods. The English temperature curve before AD 1100 is rather uncertain (H. H. Lamb, personal communication) and cannot be used for control purposes. Ice-isotope investigations, however, indicate that colder conditions prevailed around AD 1000, as do the Danish bogs. Younger variations from dark to light peat formation thus apparently reflect climatic changes and the same is probably true for older peat variations. Figure 4 also shows that there may be little (probably less than 50 yr) or no time lag between dated humification changes and the beginning of long term climatic trends.

Following the most recent humification change during
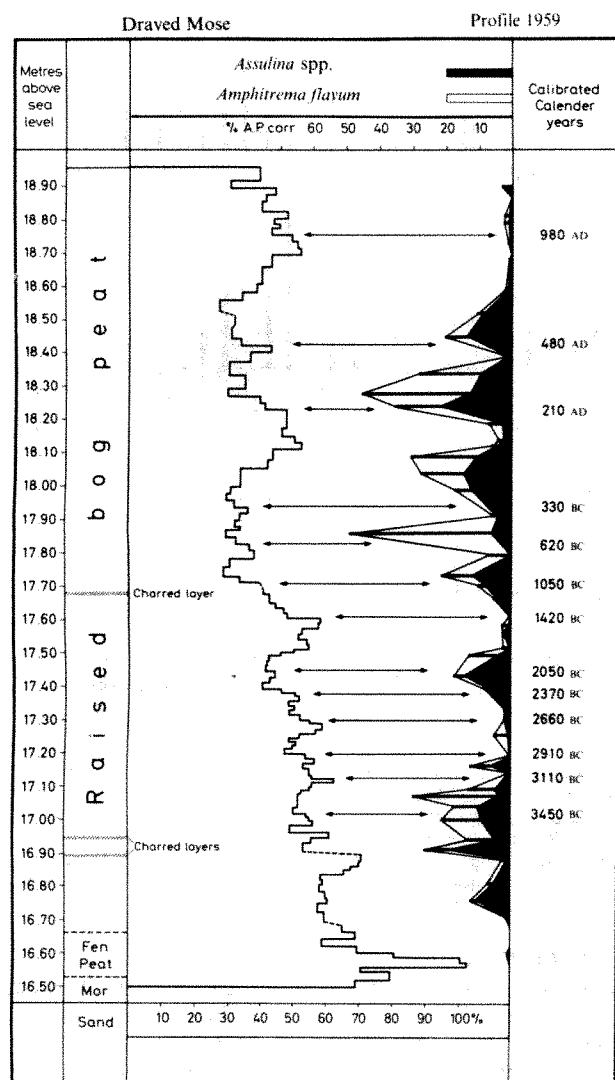
Fig. 2 Degree of humification and content of two rhizopod genera in a vertical peat column in hollow peat from Draved Mose (S₁, Fig. 1). Arrows, Levels assumed to reflect climatic shifts.

Fig. 3 Record of climatically conditioned shifts from dark to light peat formation, derived from sections in 5 Danish raised bogs. Dotted lines, dates of changes reflected in Draved Mose; full lines, results from other bogs.

the first part of the sixteenth century, another change should have occurred during part of the eighteenth century. It may indeed be noticed that a decrease in temperature occurred at that time on the England temperature curve (Fig. 4). That cooling trend was distinctly weaker than preceding ones in the thirteenth and sixteenth centuries, and was perhaps too small to leave traces in the Danish bogs.

Statistical calculations

A statistical test was made to check the significance of the apparent periodicity of humification patterns, and also to determine more accurately the length of the period.

Groups of dates in Fig. 3 differing by less than the statistical uncertainty of the datings have been considered as reflecting the same climatic events, dated to an arithmetic mean. This procedure reduces the number of events under consideration from 39 to 17, each having a smaller statistical dating uncertainty than if the dates were considered independently.

Series of dates were established with constant time intervals $T=250$, 260 and 270 yr, respectively. For each value of T the phase of the series was chosen so as to minimise the deviation of the observed dates of events from the expected dates; for example, AD 1520 was used as the youngest date in the series $T=260$ yr. It was assessed

Table 1 Distribution tests of dates indicating climatic changes (see Figs 3 and 5*b*) from three cyclic periods*

	250 yr	260 yr	270 yr
Uniform distribution test	Accepted	Refused	Moderately accepted
Significance level	20%	≪1%	9%
Normal distribution test		Accepted	Moderately accepted
Significance level		77%	5%
Minimised standard deviation		±41 yr	±74 yr

* Dates lying close together in time are supposed to reflect the same variation in climate and are only taken as a single event, dated to the arithmetic mean of the dates, which minimises standard deviations.

whether the deviations from the three periods were uniformly distributed over the entire interval. Table 1 shows that this seems to be the case for $T=250$ years, indicating no periodicity of that length.

The deviations in the two other series may show a systematic variation, and the probability that they are normally distributed has been tested (Table 1). Both the normal distribution test and the standard deviation test identify $T=260$ yr as a significant periodicity in the original series of events (Fig. 5). The standard deviations on the calibrated radiocarbon dates were not included in the calculations, but it seems likely that the calculated standard deviation in Table 1 is of the same order as, or slightly smaller than, the standard deviation of the individual calendar dates. This may indicate that in most cases the standard deviation is attributable to the uncertainty in the dating method. I thus propose that a major long term fluctuation in climate has had a rather distinct cyclicity with a periodicity of about 260 yr during the past 5,500 yr in north-western Europe.

Fig. 4 Relationship between long term trend of temperatures in central England, shown as 50-yr means, and dates of climatic shifts indicated by bog investigations (full lines). Temperature curve based on instrumental observations back to AD 1700, and before that time on historical records; the oldest part is rather uncertain (H. H. Lamb, personal communication). The curve has been smoothed by a 200-yr low pass digital filter (see refs 5 and 6). An expected change in peat formation around AD 1780 was not found in the bog profiles and is shown as a dotted line.
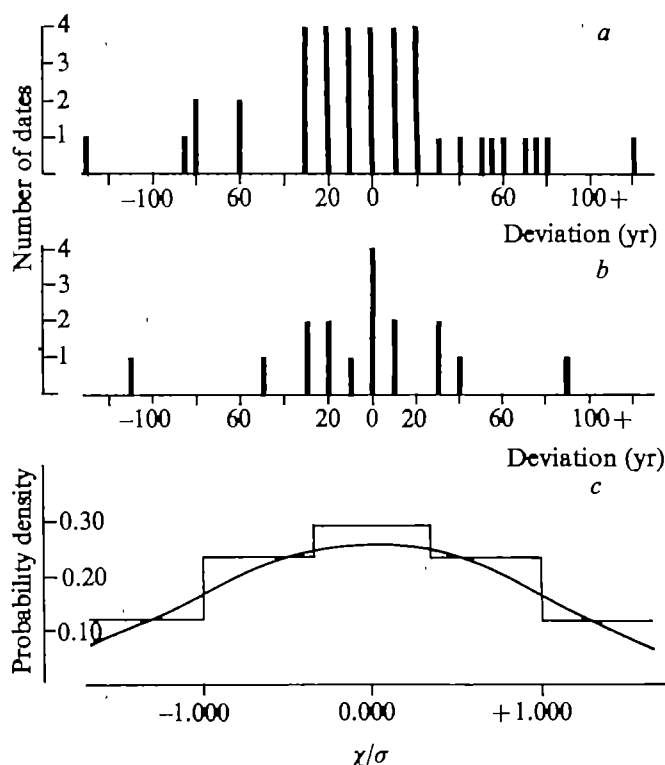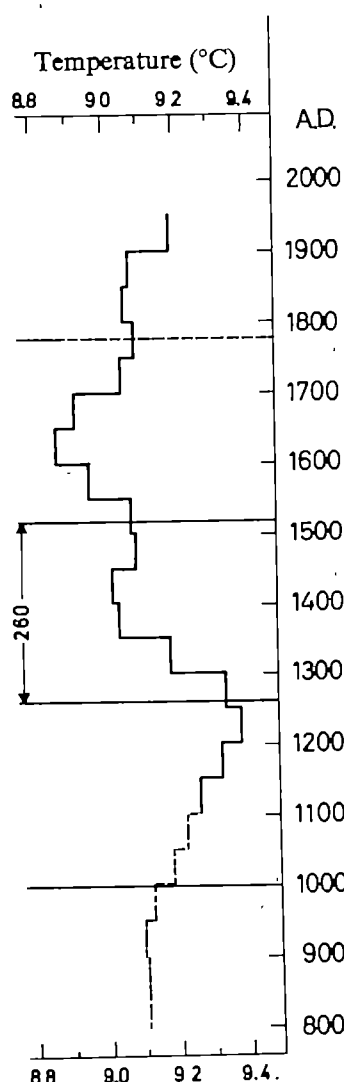
Fig. 5 *a*, Deviation (yr) from a climatic periodicity of 260 yr, beginning in AD 1520 (see Fig. 3). *b*, The same calculations based on the assumption that shifts dated to almost the same time (using ^{14}C analysis) are strictly synchronous and dated to the mean of the dates. *c*, The actual distribution of the 17 deviation dates from *b* compared with the normal distribution curve having the same mean and variance.

Future climate

To demonstrate periodicity, a time span of generally more than 10 times the length of the individual periods concerned must be considered. In this investigation the time interval has been extended to about 5,500 yr, or more than 20 times the length of the period in question. It is therefore justifiable to use the observed periodicity to predict future climatic trends. This suggests that a general trend towards decreasing mean temperatures and/or wetter conditions may begin in the last part of this century or in the first part of the twenty-first century in north-western Europe.

Predictions of long term trends in climate have previously been based on ice-isotope[7] and tree-ring[8,9] data from the past millennium, and even though these predictions essentially agree, caution must be applied in interpretation particularly as a physical explanation for climatic changes is still lacking. Pollution of the atmosphere may turn out to be a dominating factor in the future[10]. Perhaps resulting in climatic trends different from projected natural trends.

On the other hand, the Danish bog data could contribute to the design of an atmospheric model sufficiently reliable to provide some hints as to the future trends.

1 Sernander, R., *Geol. For. Stockh Forh*, 31, 423–448 (1909)
2 Aaby, B, and Tauber, H, *Boreas*, 4, 1–17 (1974)
3 Damon, P. E., Long, A, and Walihk, E. J., *Proc int Datng Conf, New Zealand*, A28–A43 (1973).
4 Aaby, B, *Geol Surv. Dan Yb 1974*, 91–104 (1975).
5 Clymo, R. S, *J. Ecol.*, 53, 747–758 (1965)
6 Lamb, H. H., *The Changing Climate*, Section 7 (Methuen, London, 1966)
7 Dansgaard, W, *et al.*, *Nature*, 255, 24–28 (1975).
8 Siren, G., *Commun Inst. Forest Fenn.*, 54, 1–64 (1961).
9 Sirén, G, *New Scient.*, 19, 18–20 (1963)
10 Jorgensen, S. E., and Meyer, H., (in the press).

Temperature (°C)

Complementary base pairing and the origin of substitution mutations

Michael D. Topal & Jacques R. Fresco

Department of Biochemical Sciences, Princeton University, Princeton, New Jersey 08540

On the basis of chemical considerations and model building, the Watson–Crick concept of complementary base pairing is extended to a wider range of DNA pairs than A–T and G–C (including A–C, G–T, A–A, G–G and G–A) by invoking imino or enol tautomers (or protonated species) and syn isomers. The virtual absence of these additional base pairs from DNA is explained in terms of the low frequency with which these unfavoured forms occur and the two-step mechanism of DNA synthesis, whereby residues are first incorporated by the DNA polymerase and then checked. This base-pairing hypothesis is used to explain the origin, nature and level of spontaneous substitution mutations, their enhancement by base analogues, and the unique effects of certain mutator alleles.

ALTHOUGH the bi-helical complementary base-paired structure of DNA implies mispairing events as the principal means for generating substitution mutations, a general mechanism for mispairing has not emerged. Watson and Crick did suggest as part of their formulation of replication that spontaneous mutations might be due to the occasional occurrence of a pu–pyr pair with a base in one of its less likely tautomeric forms[1]. But while plausible, this suggestion fails to explain the observed frequencies of spontaneous substitution mutations, 10^{-9} to 10^{-12} per base pair synthesised[2], simply from the estimated frequencies of unfavoured tautomers in solution, 10^{-4} to 10^{-5} (refs 3–6). Moreover, while pu–pyr mispairing events involving minor base tautomers have been conceived as a qualitative rationale for the origin of transitions, a comparable set of mispairs leading to transversions has not been proposed. A mispairing rationale for the mechanism of many chemical mutagens is also wanting.

We describe here a general base-pairing hypothesis that explains both qualitative and quantitative features of substitution mutations. This hypothesis is based on well recognised chemical features and isomeric equilibria of nucleotides and the notion that the expression of these equilibria accounts for the level at which mispairing events occur. The hypothesis depends on two principal assumptions. The first is that there is a wider set of complementary base pairs than the two proposed by Watson and Crick[7], A–T and G–C, that are compatible with the steric constraints of a regular DNA helix (Fig. 1). The non-Watson–Crick complementary pairs constitute mispairs of two types, pu–pyr, which can give rise to transitions, and pu–pu, which can lead to transversions. The frequency with which such mispairs occur is due mainly to the manifestation of the equilibrium constants for the isomerisation processes required for their formation, that is, keto–enol and amino–imino tautomerisation of the base moieties for the pu–pyr mispairs and both such tautomerisation and anti-syn isomerisation about the glycosyl bond of the nucleoside moiety[8,9] for the pu–pu mispairs. These two classes of mispairs can give rise by single mispairing events to all conceivable substitution mutations involving the commonly occurring nucleotides in DNA (Table 1).

The second major assumption is that there are two opportunities to express the relevant isomeric equilibria involved in particular mispairing events during the process of adding a base pair, first during catalytic incorporation of the new base on the growing strand, and again during a checking step. This results in the relevant equilibria limiting the frequency of mispairing to the level seen in biological systems.

Using these two assumptions, we are able to explain a wide range of mutagenic phenomena, including: how mispairing for each conceivable type of substitution mutation can lead to the frequencies observed *in vivo*; the rate observed for transitions and transversions[10]; the origin of the unidirectional effect of the Treffer's mutator gene on the A–T→C–G transversion[11,12]; and the specificity of the base analogues 5-bromouracil (5-BrU) and 2-aminopurine (2-AP) in inducing transitions but not transversions[13,14]. In what follows, we present the essential features of the hypothesis and the way in which it explains these basic phenomena of mutagenesis. A preliminary report of this work has been made[15].

Chemical plausibility of proposed base pairs

Each of the proposed mispairs with the H-bonding schemes indicated in Fig. 1 was studied by model building and shown to be sterically compatible with a Watson–Crick helix. Moreover, base oppositions corresponding to each of these mispairs have been shown experimentally to take up intrahelical conformations in long polynucleotide helices dominated by Watson–Crick pairs (ref. 16 and M.D.T., H. Chen and J. R. F., unpublished). By contrast, those base oppositions not invoked for mispairing events (pyr–pyr), were experimentally found to take up only extrahelical conformations in such polynucleotide helices (refs 16, 17 and J. Balcerski, M.D.T., and J.R.F., unpublished). Moreover, these oppositions could not be satisfactorily built into a regular DNA helix as base pairs. The proposed pu–pyr mispairs conform precisely to Watson–Crick helia geometry, whereas the proposed pu–pu mispairs require a slight ($\sim$ 9°) distortion of the glycosyl bond angle of the base in the syn configuration without any effect on the glycosyl bond separation distance. Distances between H-bonded atoms are standard. Finally, there is much chemical and crystallographic evidence for the occurrence in certain small molecules of the invoked unusual base tautomers[3–6,18] or nucleoside isomers (reviewed in ref. 19).

Mispairing and DNA synthesis

DNA synthesis seems generally to be a two-step process, involving an incorporation reaction, followed by a checking reaction[20]. In the cases of bacteriophage T4 (ref. 21) and *Escherichia coli*[22] at least, both these catalytic activities reside in a single polypeptide. We assume that the synthetic element of these polymerases does not directly sense the incoming base; instead, the whole DNA synthetic apparatus, consisting of several different proteins in addition to the polymerase[23,24] confers a fixed orientation on the immediate template residue and fixes the orientation and the glycosyl bond separation distance at the growing strand end. Covalent incorporation of a complementary nucleotide is seen as involving nucleoside triphosphate binding to both the polymerase on the synthetic apparatus and the geometrically constrained template residue in a specific way that is determined by H-bonded base pairing.

This binding of monomer is followed by catalysis of phospho-diester bond formation as a natural consequence of complementary base pairing, since only then is the polymerase catalytic site and the 3'-OH of the growing strand in position to attack the α–β pyrophosphate bond of the nucleoside triphosphate. The polymerase-catalysed reaction therefore incorporates any nucleotide when its base is in an isomeric form that allows a complementary pair and discriminates effectively against all others. This mechanism presupposes that once productive pairing or mispairing occurs, isomeric re-equilibration is prevented because the synthetic apparatus, perhaps by inhibiting the breathing of base pairs or by excluding water, makes the H-bonding donor and acceptor sites of the bases inaccessible to the water required for tautomerism. Thus, not only the Watson–Crick base pairs, A–T and G–C, will be incorporated, but also the complementary mispairs, which after another round of replication show themselves as substitution mutations (Table 1).

At this stage, mispairs involving only unfavoured tautomers will have occurred at a level reflecting the sum of the frequencies of unfavoured triphosphate tautomers and of such tautomers in the template residue (since equivalent A–C and G–T pairs occur when either the template or the monomer residues are in the unfavoured tautomeric form). By contrast, mispairing involving syn isomers will occur at a level reflecting the product of the frequencies of the required unfavoured tautomer in the template strand and the syn nucleoside triphosphate. Syn isomers are forbidden in the template strand since, at the outset, those residues are anti in a regular DNA helix[7], and both the stacking in the single-stranded template segment and the rigidity imposed by the synthetic apparatus must create a large barrier to rotation about the glycosyl bond. Thus, those mis-

pairings that lead to transitions will have been introduced to a level of 10^{-4} to 10^{-5} (the frequency of unfavoured tautomers), while those leading to transversions will have been introduced to a level of 10^{-4} to 10^{-5} times ~ 10^{-1} if the syn base is G[8,9] or times ~ 5×10^{-2} if it is A[8,9]. Thus, the fidelity in this synthetic step is greater for transversions than for transitions. These values are, of course, far greater than the biological levels of mutation.

Mispairing and the checking step

Once incorporated, a base pair must be checked to increase the level of fidelity by some mechanism that can remove mispairs with great efficiency. A mechanism by which this might be achieved is suggested by the occurrence of a 3'–5' exonuclease activity in the polymerase moieties of bacteriophage T4 and *E. coli*[21,22], which apparently excises unpaired residues at the growing strand terminus[25]. In higher species there are indications that this activity may be relegated to a separate molecular entity[26].

It is reasonable to assume that the isomeric equilibrium constants are about the same for the residues of the new terminal base pair as for the triphosphates and for the template residues. In the interval between release of the new base pair by the polymerase and the coming into position of the 3'–5' exonuclease moiety, there is renewed opportunity for breathing and accessibility of solvent. Therefore, we expect for transitions the same isomeric equilibria manifest in the synthetic step to be expressed once again during the checking reaction. The checking of transition mispairs should thus lead to a squaring of the fidelity level of the synthetic step, that is, a total fidelity of 10^{-8} to 10^{-10}. For transversions, however, the situation is more complex. For this class of substitutions, also, tautomeric re-

Fig. 1 The complementary base pairs that can be formed from the DNA bases A, G, T, and C are illustrated. The isomeric forms of the bases that predominate in 'physiological' conditions allow only the two Watson–Crick pairs (*a, b*). Tautomerism involving only one proton transfer gives rise, however, to two equivalent pairing schemes each for the pu–pyr pairs, A–C (*c, e*) and G–T (*d, f*) and one scheme each for the pu–pu pairs A–A$_{syn}$ (*g*), G–G$_{syn}$ (*h*), A–G$_{syn}$ (*i*) and G–A$_{syn}$ (*j*). There is no way of constructing a complementary pyr–pyr pair. Note also that the glycosyl bonds of pairs *a–f* are twofold symmetry related regardless of which strand is the fixed template, whereas pairs *g–j* are complementary only when the anti residue is on the template. The enol form of either G or T leads to a G–T pair with three H bonds (*d, f*), whereas the imino form of either A or C makes for A–C pairs with two H bonds (*c, e*). The imino form of A also makes possible A–G (*i*) and A–A (*g*) pairs with two H bonds, provided that the second base G and A, respectively, is in the syn configuration. It should be noted that the A–G pair shown in (*i*) can also be achieved by protonation of the A residue at N1 without tautomerism to the imino form at approximately the same cost in energy. Base pairs involving two H bonds are possible between a tautomerised G and either A$_{syn}$ or G$_{syn}$ (*j, h*). Here, the conventional enol tautomer of G may give rise to a close contact between a 2-amino hydrogen of G and the C8 hydrogen of the opposing A$_{syn}$ or G$_{syn}$, or else this C8 hydrogen may be restricted from H bonding to water. Therefore, a chemically possible minor tautomeric form of G is used instead, an enol, imino tautomer formed by transfer of a 2-amino hydrogen to the 6-oxygen to achieve G–G$_{syn}$ and G–A$_{syn}$.

Purine–Pyrimidine | Purine–Purine

A ≡ T *a* | *b* G ≡ C | A$_{imino}$ = A$_{syn}$ *g* | *h* G$_{enol, imino}$ = G$_{syn}$

A = C$_{imino}$ *c* | *d* G ≡ T$_{enol}$ | A$_{imino}$ = G$_{syn}$ *i* | *j* G$_{enol, imino}$ = A$_{syn}$

A$_{imino}$ = C *e* | *f* G$_{enol}$ ≡ T
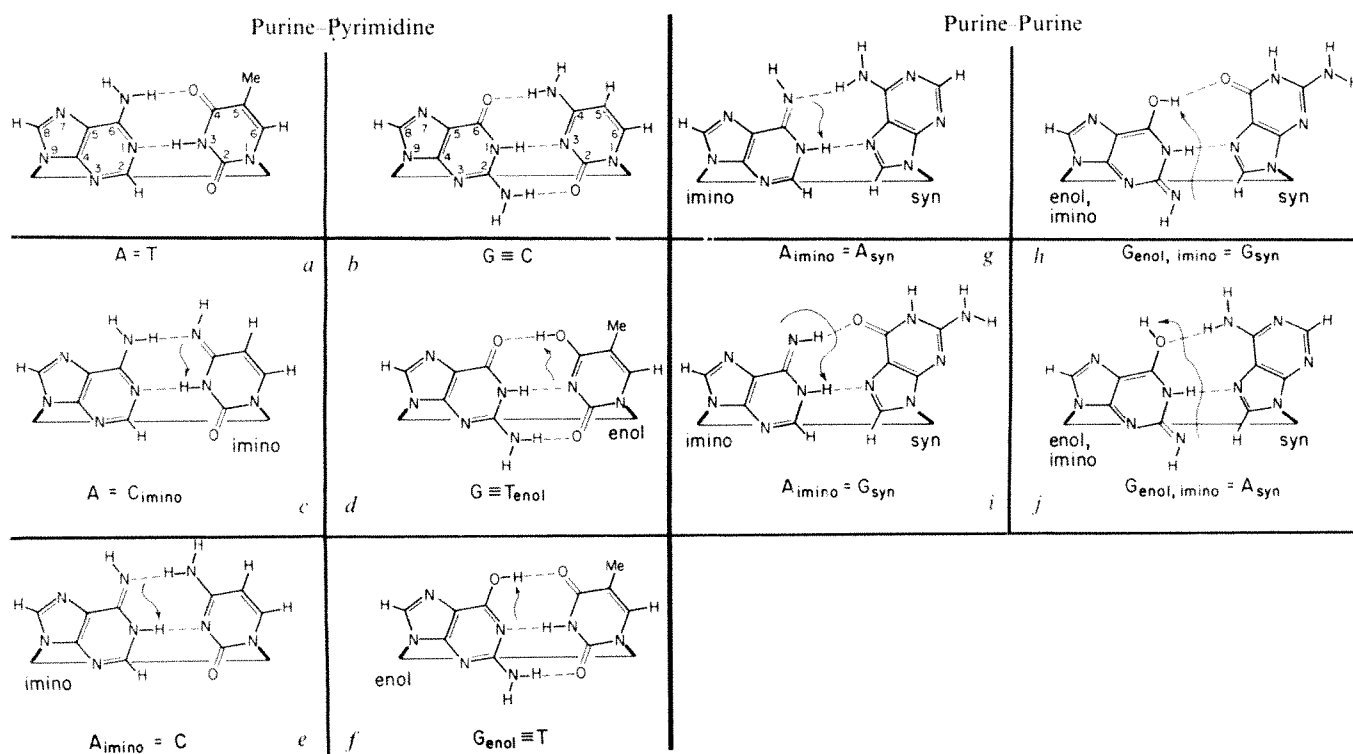
Table 1 Pathways for substitution mutations
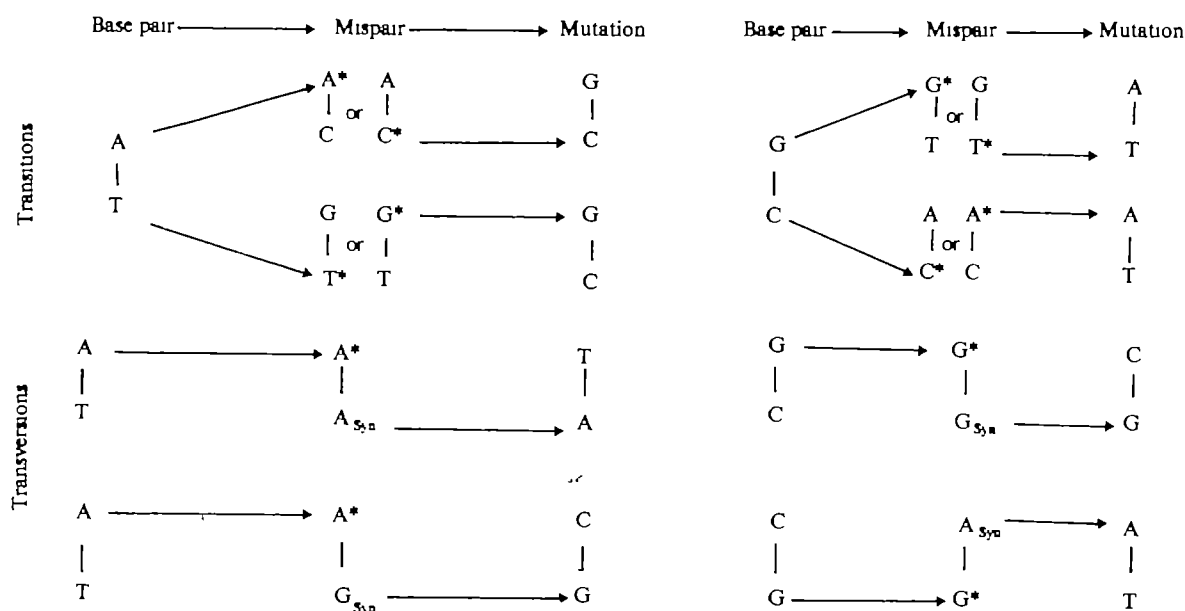
Pathways are shown whereby every conceivable transition and transversion can occur by single mispairing events that each involve a non-Watson–Crick complementary base pair. Note that whereas transitions can occur with each strand of a base pair, transversions are viewed as occurring only on the strand with the purine template residue since pyr–pyr mispairs are not believed to occur with any significant frequency.
*, Base in a minor tautomeric form; syn, the unfavoured configuration about the glycosyl bond of that residue (see Fig. 1). Arrows indicate rounds of replication.

equilibration of template residues is expected in the checking step; but re-equilibration of the syn configuration of the terminal residue is unlikely because the activation energy for rotation about the glycosyl bond of an incorporated residue stacked on its neighbour should be much larger than it is in the entering triphosphate Consequently, the exonuclease step should enhance fidelity only by the frequency of the minor tautomer in the template residue, that is, by another factor of 10^{-4} to 10^{-4}. Therefore, the total fidelity for transversions will be less than the square of that achieved in the first step, amounting to 10^{-4} to 5×10^{-11}, which is nevertheless greater than that for transitions. It follows that in systems lacking the exonuclease function, substitution mutation rates will be as expected from the first step alone.

Evaluation of the hypothesis

In evaluating the hypothesis, we have considered the literature on nucleic acid chemistry, on the specificity of *in vitro* and *in vivo* nucleic acid biosynthesis and on the mechanism of mutagenesis. We have not found any experimental observations that could not be interpreted in terms of our two principal assumptions. Here, we describe three important types of observations on substitution mutations that illustrate the consistency between hypothesis and observation.

(1) One of the major probes of the mechanisms of mutagenesis involves studies with compounds that induce or enhance the level of spontaneous mutations From the work of Benzer and Freese, in particular, it has been established that certain base analogues are highly specific in inducing transitions only[13,17]. For example, pyrimidines such as 5-BrU, which exhibit a greater tendency to enolise than does their natural counterpart, thymine, behave this way. This selective effect is exactly as expected from our hypothesis since transitions are mediated only by pu–pyr mispairs, whereas transversion mispairs do not involve pyrimidine residues

2-AP is another base analogue that induces only transitions[13,14]. The inability of this purine to induce transversions is of particular interest in view of our proposal that transversions are brought about only by pu–pu mispairs.

The way in which this analogue induces transitions is not difficult to rationalise. Although lacking a substituent on C6, it can interact "normally" with T on the template to form a complementary pair with two H bonds. By tautomerism to the imino form, it can "mispair" with C also by way of two H bonds. The alternative mispairing scheme with C, involving the amino tautomer of 2-AP seems less likely because it allows only one H bond, and the H-bonding tendencies of the internal N1 of the purine and N3 of the pyrimidine would be left unsatisfied.

The lack of potential for pu–pu mispairs involving 2-AP can

Fig. 2 Potential of 2-AP for pu–pu pairing. The inability of 2-AP to form transversion mispairs is illustrated. In the syn configuration, 2-AP can form no H bonds with A (*a*), and whereas a pair with one H bond might be formed with G, it results in a bad Van der Waals' contact (*b*). When in the anti configuration, 2-AP can make but one H bond with either A (*c*) or G (*d*), and then only by tautomerising to the less favoured imino form.
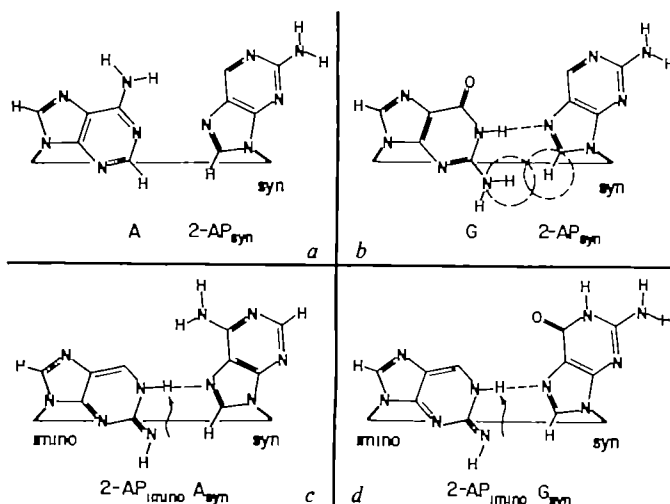
Table 2 Comparison of observed frequencies and calculated mutation rates

	Base pair substitution	Mutator allele*	Observed† reversion frequencies	Total	=	Predicted reversion rates‡ Synthetic step	×	Checking step	×	No. of pathways
Transitions	AT→GC	+	5×10^{-9}–4×10^{-10}	5×10^{-9}	=	(5×10^{-5})	×	(5×10^{-5})	×	2
		D5	$(1–2) \times 10^{-5}$	1×10^{-4}	=	(5×10^{-5})	×	(1)	×	2
	GC→AT	+	$(3–9) \times 10^{-10}$	5×10^{-9}	=	(5×10^{-5})	×	(5×10^{-5})	×	2
		D5	$(2–4) \times 10^{-5}$	1×10^{-4}	=	(5×10^{-5})	×	(1)	×	2
Transversions	GC→CG	+	8.6×10^{-10}	2.5×10^{-10}	=	$(5 \times 10^{-5})(1 \times 10^{-1})$	×	(5×10^{-5})	×	1
		D5	2.9×10^{-6}	5×10^{-6}	=	$(5 \times 10^{-5})(1 \times 10^{-1})$	×	(1)	×	1
	AT→TA	+	3.9×10^{-9}	1.2×10^{-10}	=	$(5 \times 10^{-5})(5 \times 10^{-2})$	×	(5×10^{-5})	×	1
		D5	$(2–3) \times 10^{-6}$	2.5×10^{-6}	=	$(5 \times 10^{-5})(5 \times 10^{-2})$	×	(1)	×	1
	CG→AT	+	$(3–5) \times 10^{-9}$	1.2×10^{-10}	=	$(5 \times 10^{-5})(5 \times 10^{-2})$	×	(5×10^{-5})	×	1
		D5	3×10^{-5}–2×10^{-6}	2.5×10^{-6}	=	$(5 \times 10^{-5})(5 \times 10^{-2})$	×	(1)	×	1
	AT→CG§	+	2×10^{-8}–2×10^{-9}	2.5×10^{-10}	=	$(5 \times 10^{-5})(1 \times 10^{-1})$	×	(5×10^{-5})	×	1
		D5	5×10^{-6}–3.1×10^{-7}	5×10^{-6}	=	$(5 \times 10^{-5})(1 \times 10^{-1})$	×	(1)	×	1
	AT→CG§	+	$(6–9) \times 10^{-10}$	2.5×10^{-10}	=	$(5 \times 10^{-5})(1 \times 10^{-1})$	×	(5×10^{-5})	×	1
		D5	1×10^{-7}–5×10^{-8}	5×10^{-6}	=	$(5 \times 10^{-5})(1 \times 10^{-1})$	×	(1)	×	1

Comparison of frequency of different substitution reversions in the tryptophan synthetase *A* gene of *E. coli* (data taken from Fowler, *et al.*[10]) with theoretical mutation rates calculated from characteristics of proposed non-Watson–Crick complementary base-pair intermediates.

*+ wild type; *D5*, conditional mutator gene, to which we assign the phenotype, exonuclease⁻.

†Ranges are given to encompass several sets of data obtained for the same mutant by Fowler *et al.*[10].

‡For the purposes of these calculations, a value of 1×10^{-1} is taken for the equilibrium frequency of G_{syn} and 5×10^{-2} for that of A_{syn}. The same equilibrium frequency, 5×10^{-5} is assumed for all unfavoured tautomers. If this were actually the case, these calculations would have to take into account that the A–C and G–T pairs each could occur with either the template or the monomer residues in the unfavoured tautomeric form; then, the calculated rates of transitions would have to be raised by a statistical factor of four (two for each step). It is much more likely, however, that the tautomeric equilibria differ for the two members of each pair and differences as small as two–threefold effectively reduce this factor to one. With only one unfavoured tautomer in pu–pu pairs, this factor is also one for transversions.

§These represent the same type of transversion at two different sites within the same gene.

be seen (Fig. 2) to arise from its lack of any substituent on C6. For incorporation opposite a purine, 2-AP must enter in the syn configuration, in which case it can form no H bonds with A (Fig. 2*a*). Against G, the possibility of only one H bond as well as the potential close contact between the C2 amino of the template residue and the C8 hydrogen of 2-AP, which can also interfere with solvation of the C2 amino group, prohibits that pairing as well (Fig. 2*b*). Further, once incorporated, the amino tautomer of 2-AP in the anti configuration of the template strand cannot form any H bonds with either A_{syn} or G_{syn} triphosphates, while the imino tautomer can form pairs which will allow but one H bond (Fig. 2*c* and *d*), which would seem insufficient. Purine analogues that favour syn configurations and have appropriate substituents are, however, expected to induce transversions; and purine analogues of that type which also have a more favourable capacity to tautomerise than A or G should enhance the rate of both transitions and transversions. It may similarly be possible to explain the mutagenic effects of reactive substances that modify bases of intact DNA, thereby altering their capacity either to tautomerise (for example, hydroxylamine[28,29] and Cu^{2+} (ref. 30)) or to be fixed in a syn configuration. Since many such reactive compounds are also carcinogens[31], a correlation is suggested.

(2) Among mutator genes, there is a striking example discovered by Treffers[11] of an allele in *E. coli* that selectively enhances the rate of the A–T→C–G transversion but not the reverse transversion[12]. A consideration of the pu–pu mispairs for these two transversions (Fig. 1*i,j*) and the restriction that the syn residue of those pairs be the entering one, suggests how the unidirectional effect of this mutator comes about. Whereas the forward transversion must be mediated by A_{imino}–G_{syn}, the reversion can only be mediated by $G_{enol,imino}$–A_{syn} and not by the mispair by which it is formed (Table 1). The physiological mechanism by which this mutator is expressed can be imagined to involve a gene coding for some component of the replicative apparatus that allows incorporation of G_{syn} but not A_{syn}. (This would also result in parallel discrimination between the A–A_{syn} and the G–G_{syn} mispairs.) Obviously, a mutator with this kind of discrimination could not affect the rate of transitions because only anti isomers participate in transition

mispairs. The discriminatory characteristics of the Treffers' mutator and others[32,33] clearly correlate with the proposed features that distinguish the mispairs that give rise to transitions from those that lead to transversions.

(3) The rates of transitions and transversions are also consistent with our proposed base-pairing schemes. Specific data against which to evaluate the mechanisms proposed for substitution mutations are provided by a study[10] of spontaneous substitution mutation frequencies for particular sites on the tryptophan synthetase A gene of *E. coli*[14], for which the base-pair changes are known. Those frequencies, determined both in the absence and presence of a mutator gene, *mutD5*, are shown in Table 2. Presented for comparison are the rates we calculate for each transition and transversion from the relevant isomeric equilibria for the proposed complementary non-Watson–Crick mispairs. (For this comparison, it is assumed that *in vivo* mutation rates and the observed frequencies they generate are equivalent. In fact, the observed values should be high, but by no more than a factor of five[32].) Because of the availability of two mispairing routes to both transitions, but only one to each transversion (see Table 1) the rates calculated for the transitions (only) contain a factor of two. In the case of *mutD5* we make the explicit assumption that this mutation, though not in the exonuclease gene[34], has inactivated the checking function. The obvious correlation that this assumption allows us to make between the observed frequencies and the rates calculated, reinforces our confidence in the assignment of the mutD5 phenotype. Thus, it can be seen that the agreement between the observed reversion frequencies for the two transitions examined in the presence of *mutD5* and our estimate for the synthetic step alone is satisfactory; so is that between the frequencies measured in the absence of *mutD5* and our calculated rates. Furthermore, the experimental data agree with our expectation that the total fidelity must approximately equal the square of the frequencies of the unfavourable tautomers only for transitions, where the complementary intermediates are A–C and G–T pairs.

The complexity postulated for transversions is also evident. As demanded by our mechanism and consistent with separating the *mutD5* frequencies from the reversion frequencies observed

in the absence of this mutator, greater fidelity can be seen to be achieved for transversions in the synthetic step than in the checking step. This is because syn isomers should have no role in the checking step. The *mutD5* frequencies are also consistent with our expectation that the synthetic step achieves greater fidelity for transversions than for transitions because of the involvement of syn isomers as well as unfavourable tautomers in transversion mispairs. Indeed, agreement between the observed frequencies in the presence of *mutD5* and the rates calculated for the synthetic step alone for all transversions is quite close, considering that the syn–anti equilibria must be taken into account. Nevertheless, it is apparent that the overall fidelity observed for transversions is somewhat less than we expect. Conceivably, this is because of a lower efficiency in removal of syn than anti residues by the exonuclease. Such a differential effect towards syn compared with anti residues has been observed for other exonucleases[24].

The agreement in both trends and levels between observed and calculated values for the various substitution mutations is encouraging. Nevertheless, we recognise that frequencies are being compared with rates and that the calculations do not take into account such modulating effects as nearest neighbour interactions (which could account for the frequency differences observed in the two cases of the same transversion A–T→C–G at two different loci shown in Table 2) and differences in the relative stability of the various mispairs. Since the calculated rates follow the observed trend, these modulating effects must be small relative to the major factors that we have introduced in the calculations.

The hypothesis we have presented here concerning a new set of complementary base pairs and their relevance to mutagenesis is chemically reasonable and consistent with a large body of biochemical and genetic facts. We therefore believe that the main features of these concepts are correct, though the details may require modification. Experiments explicitly designed to test particular features of the hypothesis are needed to assure their validity. We are initiating some, and we hope that others will be prompted to design additional tests.

The wider concept of complementary base pairing that has been introduced here has relevance to other problems in molecular biology. In the accompanying paper[36], we address the question of their role in protein synthesis. Later reports will deal with their relevance in other contexts.

1 Watson, J D , and Crick, F. H. C., *Nature*, 171, 964–967 (1953).
2 Drake, J. W , *Nature*, 221, 1132 (1969).
3 Tucker, G., and Irvin, J., *J. Am. chem Soc.*, 73, 1923–1929 (1951).
4 Kenner, G. W., Reese, C. B , and Todd, A R., *J. chem. Soc*, 855–859 (1955).
5 Katritzky, A. R., and Warring, A J , *J. chem. Soc* , 1540–1544 (1962).
6 Wolfenden, R. V., *J. molec. Biol* , 40, 307–310 (1969).
7 Watson, J. D., and Crick, F. H C., *Nature*, 171, 737–738 (1953).
8 Jordan, F , and Pullman, B., *Theor chim Acta*, 9, 242–252 (1968)
9 Tinoco, L , Jr, Davis, R. C., and Jaskunas, S. R., in *Molecular Association in Biology* (edit by Pullman, B.), 77–92 (Academic, New York and London, 1968).
10 Fowler, R G , Degnen, G. E., and Cox, E. C., *Molec. gen Genet.*, 133, 179–191 (1974).
11 Treffers, H. P., Spinelli, V , and Belser, N. O , *Proc. natn. Acad. Sci. U.S.A.*, 40, 1064–1071 (1954).
12 Yanofsky, C , Cox, E., C., and Horn, V. D., *Proc. natn. Acad. Sci. U.S A.*, 55, 274–281 (1966).
13 Freese, E., *Proc natn. Acad Sci. U.S.A.*, 45, 622–633 (1959).
14 Yanofsky, C , Ito, J., and Horn, V.D., *Cold Spring Harb. Symp. quant. Biol.*, 31, 151–162 (1966)
15 Fresco, J. R., and Topal, M. D., *Abstracts 100th Anniversary Meeting of Am. chem Soc., New York*, April 1976.
16 Lomant, A J , and Fresco, J. R., *Prog. Nucl. Acid. Res. molec. Biol.*, 15, 185–218 (1975)
17 Fresco, J R , and Alberts, B. M., *Proc. natn Acad Sci. U.S.A.*, 46, 311–321 (1960)
18 Daniels, M., *Proc. natn. Acad. Sci. U.S A.*, 69, 2488–2491 (1972).
19 Ts'o, P O. P., in *Basic Principles in Nucleic Acid Chemistry*, 1 (edit. by Ts'o, P O P.), 481–512 (Academic, New York, 1974)
20 Kornberg, A , *DNA Synthesis* (Freeman, San Francisco, 1974).
21 Lehman, I R , and Richardson, C C , *J. biol Chem* , 239, 233–241 (1964)
22 Goulian, M , Lucas, Z. J., and Kornberg, A , *J biol Chem.*, 243, 627–638 (1968).
23 Morris, C. F , Sinha, N K , and Alberts, B. M , *Proc. natn. Acad Sci. U S A* , 72, 4800–4804 (1975)
24 Schekman, R., Weiner, J. H., Weiner, A , and Kornberg, A., *J biol Chem* , 250, 5859–5865 (1975).
25 Brutlag, D , and Kornberg, A., *J. biol. Chem.*, 247, 241–248 (1972).
26 Lindahl, T , Gally, J A., and Edelman, G. M , *J. biol. Chem.*, 244, 5014–5019 (1969)
27 Benzer, S., and Freese, E , *Proc. natn. Acad. Sci. U S A.*, 44, 112–119 (1958).
28 Freese, E , Bautz-Freese, E., and Bautz, E., *J. molec Biol.*, 3, 133–143 (1961).
29 Champe, S P., and Benzer, S., *Proc. natn Acad. Sci. U S A.*, 48, 532–546 (1962).
30 Weed, L. L., *J Bact* , 85, 1003–1010 (1963).
31 Singer, B , *Prog Nucl. Acid. Res. molec. Biol* , 15, 219–284 (1975)
32 Drake, J W , *The Molecular Basis of Mutation*, 181–184 (Holden–Day, San Francisco, 1970.)
33 Cox, E. C., Degnen, G. E , and Scheppe, M. L , *Genetics*, 72, 551–567 (1972).
34 Cox, E. C., *Genetics Suppl.*, 73, 67–80 (1973).
35 Ward, D C , and Reich, E , *Proc natn. Acad. Sci U.S.A.*, 61, 1494–1501 (1968).
36 Topal, M D , and Fresco, J. R , *Nature*, 263, 289–293 (1976)

Base pairing and fidelity in codon–anticodon interaction

Michael D. Topal & Jacques R. Fresco

Department of Biochemical Sciences, Princeton University, Princeton, New Jersey 08540

Base pairing in codon–anticodon interaction has been investigated in order to understand the basis on which particular base pairs have been selected for or against participation at the wobble position and the basis for codon–anticodon infidelity.

STUDIES of the genetic code revealed that the Watson–Crick pairing rules for DNA[1] are only followed strictly for the first two base pairs of codon–anticodon interaction[1]. Much of the pairing at the third codon position also involves the normal Watson–Crick base pairs, A–U, and G–C, but for several codon–tRNA interactions, non-Watson–Crick pairs are clearly required. This led Crick to propose that the pair at the 3′ end of the codon can deviate (from complementary helix geometry) or "wobble" within arbitrary limits (to the glycosyl bond separation distance and dihedral angles) that are consistent with the additional pairing interactions suggested by the code.

Our reconsideration of this problem takes into account the chemical features of the bases and the stereochemistry of non-Watson–Crick complementary base pairs, described in the accompanying paper[3], that have allowed us to suggest mechanisms for mispairing in replication. We deduce that the "wobble" pairing schemes for G–U, U–G and U–I must be as proposed by Crick[4], but that an alternative base-pairing scheme for A–I is a more suitable codon–anticodon interaction. We also show that in a manner similar to the way they act in nucleic acid biosynthesis[3], complementary non-Watson–Crick base pairs contribute to the natural level of infidelity in protein synthesis. A preliminary report of this work has been made[5].

Base-pairing alternatives

The wider set of complementary base pairs described in the accompanying paper[3], utilises base tautomers (enol and imino) and nucleoside isomers (syn) that are unfavoured in aqueous solution. With U substituted for T, we must consider in the case of protein synthesis the pairs that can be formed in this way from the four characteristic bases of nucleic acids and in addition, C–I, U–I, G–I and A–I, since inosine is the first

anticodon residue of many tRNAs[6]. Figure 1 shows complementary pairing schemes (studied by model building) containing I; of these, only C–I utilises the preponderant isomers. The ability of such base oppositions to take up intrahelical arrangements within helices dominated by Watson–Crick pairs has also been experimentally demonstrated, though the actual H-bonding schemes have not been ascertained (ref. 7, and M.D.T., H. Chen and J.R.F., unpublished).

As shown in Fig. 2a–d, all the pairs covered by Crick's wobble-pairing schemes make use of the favoured forms of the residues, and the glycosyl bond separation distances are close to the 10.9 Å typical of a complementary pair, for all except A–I, for which it is 12.8 Å. Assuming that the codon constitutes a fixed template, this longer separation distance exaggerates the extent of backbone relocation required of the anticodon chain at the wobble position. In contrast, the amount of chain relocation required for those wobble pairs with the near-normal separation distance appears to be negligible for the short codon–anticodon helical array.

Additional wobble pairs—A–C (and C–A) can be built with the same stereochemistry as their G–U and U–G counterparts (Fig. 3a,b) only with a single H bond. They are unlikely to occur since this arrangement would lead to one unfulfilled H-bond

the syn residue lies in a complementary position; but when the template residue is syn, the opposite member takes up a "wobble" position. Thus, the additional pu–pu wobble pairs in Fig. 3 have their complementary inverse pairs; for example, wobble pair A_{syn}–$G_{enol,imino}$ (Fig. 3c) has its complementary pair inverse $G_{enol,imino}$–A_{syn} (Fig. 1j of accompanying paper[3]).

Base pairing in codon–anticodon interaction

From these considerations it is difficult to decide which type of pairing is most relevant to codon–anticodon interaction. Although the complementary non-Watson–Crick pairs are sterically more acceptable, the thermodynamic cost of utilising the unfavourable tautomers or isomers they require could be greater than the cost of a small helix distortion at the wobble position. We have found it possible to choose between these alternatives with the aid of the following three assumptions.

(1) The ribosome limits wobble-pairing schemes to those that minimise backbone distortion and it constrains the message as a template. Studies of coding fidelity with ribosomal mutants[8] and drugs such as streptomycin[9,10] indicate that the ribosome does impose steric constraints on codon–anticodon interaction.

(2) The ribosomal cavity for codon–anticodon interaction precludes tautomerism and protonation therein. Complementary
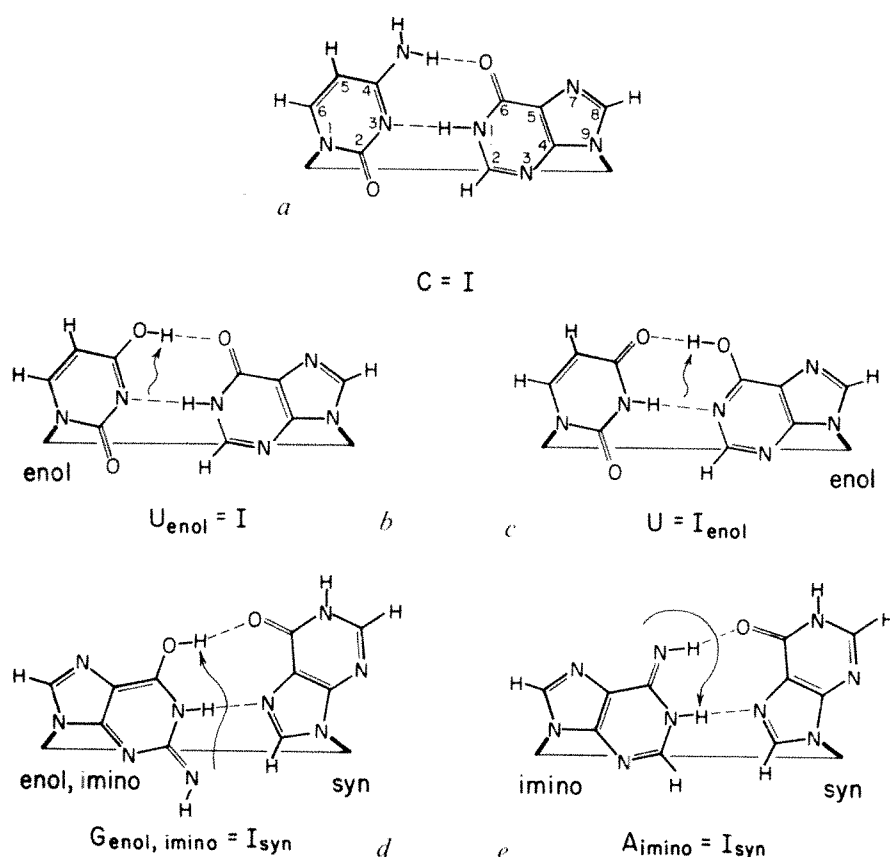
Fig. 1 Complementary base-pairing schemes involving the anticodon I residue. a–c, Pyrimidine–purine; d, e, purine–purine. The C–I pair (a) is in all respects the equivalent of a standard Watson–Crick complementary pair. Other non-Watson–Crick complementary pairs are shown in Fig. 1 of the accompanying paper[3] where, however, T is used in place of U. Note that the stereochemistry for the pu–pyr pairs that involve an unfavoured tautomer in either member conforms precisely to that of the Watson–Crick pairs, whereas that for the pu–pu pairs requires a small deviation, ∼ 9°, in the glycosyl bond angle of the anticodon residue, but no alteration in glycosyl bond separation distance. Wherever shown, arrows indicate the source of a proton that has been transferred in the tautomeric shift from a favoured to an unfavoured isomer.

acceptor site, or two H bonds made possible by protonation of the A member. Additional pu–pu wobble pairs are also conceivable (Fig. 3c–e) which make use of combinations of syn isomers and unfavoured tautomers. In particular the A–I pair shown in Fig. 3e, which makes use of a syn residue in the codon position has a near-standard glycosyl bond separation distance, instead of the much longer one proposed by Crick (Fig. 2b).

When one member of a pu–pu pair has its backbone location fixed as part of a template, an interesting asymmetry exists if one member is syn (Fig. 4). When the template residue is anti,

mispairing requires unfavoured tautomers which can occur only by a proton jump by way of water molecules proximal to the donor and acceptor atoms of the base. Binding by the ribosome of both messenger and tRNA must inhibit breathing of the codon–anticodon complex. This alone will preclude tautomerism since water is excluded between paired bases. Moreover, when tRNA and mRNA enter the cavity, bulk water that fills such space in their absence will be displaced, so that in the immediate vicinity of codon–anticodon negotiation, the amount of free water will be greatly reduced. If re-equilibra-

tion to minor tautomers is thereby precluded, miscoding due to complementary mispairs (Fig. 1 of accompanying paper[3] and Fig. 1 here) will be kept at a very low level (assumption (3)). Therefore, the specificity at all codon-anticodon positions must depend primarily on the predominant keto and amino forms of the bases.

(3) Tautomeric equilibria nevertheless play a part in codon-anticodon interaction to the extent that the message and tRNA manifest these equilibria in solution. Whereas the ribosome will prevent mismatched oppositions from becoming complementary mispairs (assumption (2)), they must also lock in those unfavoured tautomers that have matched up to form such mispairs (for

occurs at the tRNA wobble position[17] and there are 16 codons with U in the third position. The different geometry for G-U raises the question of whether this pair is also used. Comparison of the code with anticodons established in experiments with sequenced tRNAs (see Table 1 of ref. 17) suggests that G-U may not be required. Thus, Gln, Lys and Gly, with code words that have A or G in the third position (CAA and CAG for Gln, AAA and AAG for Lys, GGA and GGG for Gly), each have two corresponding tRNAs in the same species[18], one with a modified U and another with C in the wobble position. On the other hand, an unusual case has been found in which an ochre suppressor tRNA (UUA) reads the amber codon (UAG) in vivo

Fig. 2 Wobble-type or non-complementary H-bonded base pairs proposed by Crick[4] for codon-anticodon interaction. The codon or template residue is always shown on the left, and the horizontal line from the Cl' carbon of the codon residue extends to the location of the Cl' carbon of the anticodon residue if the latter were complementary. Note that only keto and amino tautomeric and neutral forms of the bases are used for the pairing schemes proposed by Crick. In each case, the twofold symmetry that relates the Cl' carbons of the two residues of a complementary pair is absent. Generally, this leads to only a small amount of backbone displacement because the glycosyl bond separation distance is maintained close to that of a complementary helix. In the case of A-I (b), however, that distance is nearly 2 Å greater, which results in excessive backbone displacement.
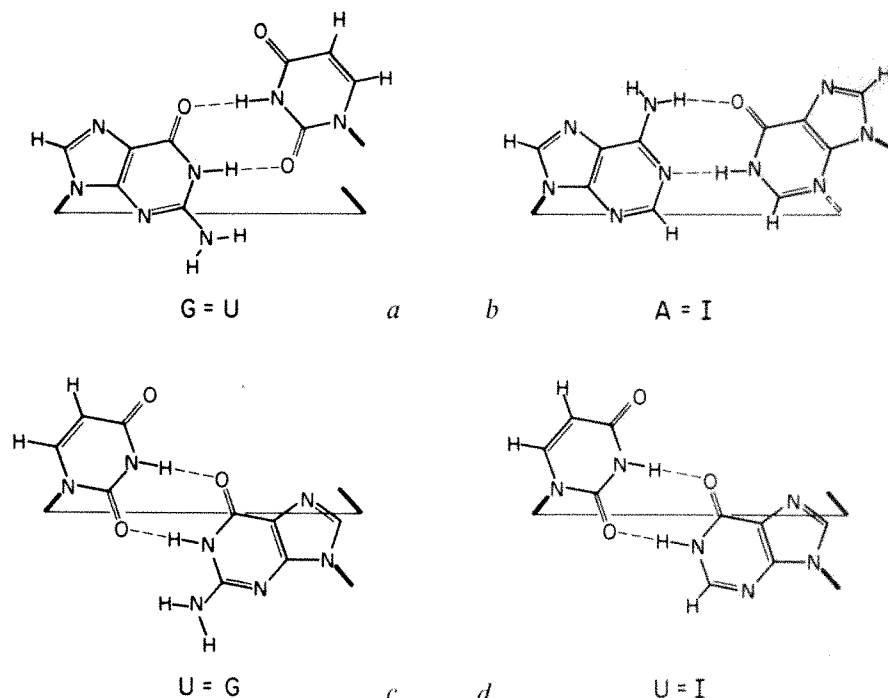
example, A_{imino} with C, Fig. 1 of ref. 3). Hence, the low frequencies of minor tautomers that both interactants must contain will be frozen by the ribosome and be expressed. This will give rise to a background in vivo level of spontaneous miscoding involving all complementary mispairs since they all involve minor tautomers (G-U, A-C, U-I, G-I, A-I, A-A, A-G, and G-G). Miscoding should therefore occur at all three positions at frequencies that depend on the relative occurrence of the isomers required in the mispairs: for pu-pyr, 10^{-4} to 10^{-5} (tautomeric frequency)[11-14]; for pu-pu, the latter values are lower, 10^{-5} to 5×10^{-7} because pu-pu pairs involve syn isomers[15,16] which occur at a frequency of $\sim 10^{-1}$ for G_{syn} and $\sim 5 \times 10^{-2}$ for A_{syn}, as well as unfavoured tautomers.

From these assumptions we can evaluate pairing schemes for those non-Watson-Crick base oppositions that prevail in the wobble position and can then deduce the nature of the mispairs that lead to infidelity. Throughout, base pairs are denoted with the codon residue on the left.

Pairing in the wobble position

U-G, U-I-. Since tautomerism is precluded, Crick's wobble-pairing schemes must prevail. With the codon as template, the glycosyl bond of G or I at the wobble anticodon position will be displaced to the same extent and in the same direction.

G-U-. The deviation of the anticodon residue for wobble G-U, though equal to that for wobble U-G, is opposite in direction of displacement of the anticodon glycosyl bond (Fig. 2a,c). U-I and U-G are required by the code since A never

very effectively[19]. In this case at least, G-U has a role even though it is sterically different from U-G.

A-I-. We suggest instead of Crick's "long" wobble pair (Fig. 2b), one in which A in the codon is syn, while I in the anticodon is anti (Fig. 3e). This arrangement, observed in a cocrystal of 9-ethyl-8-bromoadenine and O-ethyl-8-bromohypoxanthine[20], uses favoured tautomers and allows a standard glycosyl bond separation distance[21]. With codon as template, however, only wobble geometry is possible (Fig. 4). This arrangement ought to form with acceptable frequency if the 2 kcalorie mol^{-1} (refs 15, 16) cost of A_{syn} is not increased within the ribosome.

Evidence for an A-I pair in protein synthesis is not definitive. In vitro coding data [22,23] are consistent with its occurrence, but subject to artefacts due to the unnatural levels of divalent cations used[24]. Moreover, it seems that there are two tRNAs with I in their anticodons (one for Val and one for Ser) that are not needed for interaction with codons (GUA and UCA, respectively) that have A in the third position[17] since isoacceptor tRNAs exist with U in place of I. Because these isoacceptors have not yet been demonstrated in the same species, however, their bearing on the question is uncertain. Conceivably, a tRNA with I instead of U in the anticodon is required to prevent the ambiguous reading of the Met codon, AUG, by a tRNA[Ile] that normally reads AUA. It is now apparent[25,26], however, that ambiguous recognition of G instead of A is avoided by replacing U with s^2U. In any case, if A-I occurs in vivo, then the A_{syn}-I wobble arrangement would seem to be more favourable than Crick's[4].

A–G–. Although this pair occurs in long Watson–Crick helices (M. D. T., H. Chen and J. R. F., unpublished), the code forbids it at the wobble position. Crick suggested[4] that the A–G counterpart to the A–I wobble pair does not occur because the amino group of G cannot H bond even to water. If A–I has a role in protein synthesis, this presumed constraint would be equally relevant to the A_{syn}–G counterpart to the proposed A_{syn}–I pair. Moreover, it seems that a 2-amino hydrogen of G eclipses the C8 hydrogen of A_{syn}. In either case, a tautomeric shift (2-amino→6-oxygen) could remove this constraint. As noted, however, such shifts are precluded within the ribosome. If instead, A–I does not occur *in vivo*, then both A–I and A–G may be excluded by the same ribosomal constraints that prevent A from either becoming syn or pairing with a long glycosyl bond separation distance.

A–C, G–G, G–A, C–A–. These pairs also occur in long Watson–Crick helices (M.D.T., H. Chen and J.R.F., unpublished), even though they are prohibited by the code[27].

the U–G, G–U and U–I pairs there (because these pairs could not then occur with the high frequency required), while the original H-bonding schemes of the wobble hypothesis are correct except for the case of A–I. A more reasonable wobble-pairing arrangement is proposed for A–I that is consistent with the absence of the analogous A–G pair.

Mispairing in codon–anticodon interaction

We now evaluate the role of non-Watson–Crick pairs as a source of infidelity in protein synthesis. It is evident from the prohibition of wobble pairs at the first two codon positions that the ribosome effectively screens out non-complementary pairs from those positions. In normal *in vivo* conditions, however, there ought to be a background level of miscoding that reflects the frequency of residue isomers on both message and tRNA before they enter the ribosome (assumption (3)), that can give

C = A⁺   *a*   *b*   A⁺ = C

A_{syn} = $G_{enol, imino}$ *c*   *d* $G_{enol, syn}$ = I

A_{syn} = I   *e*
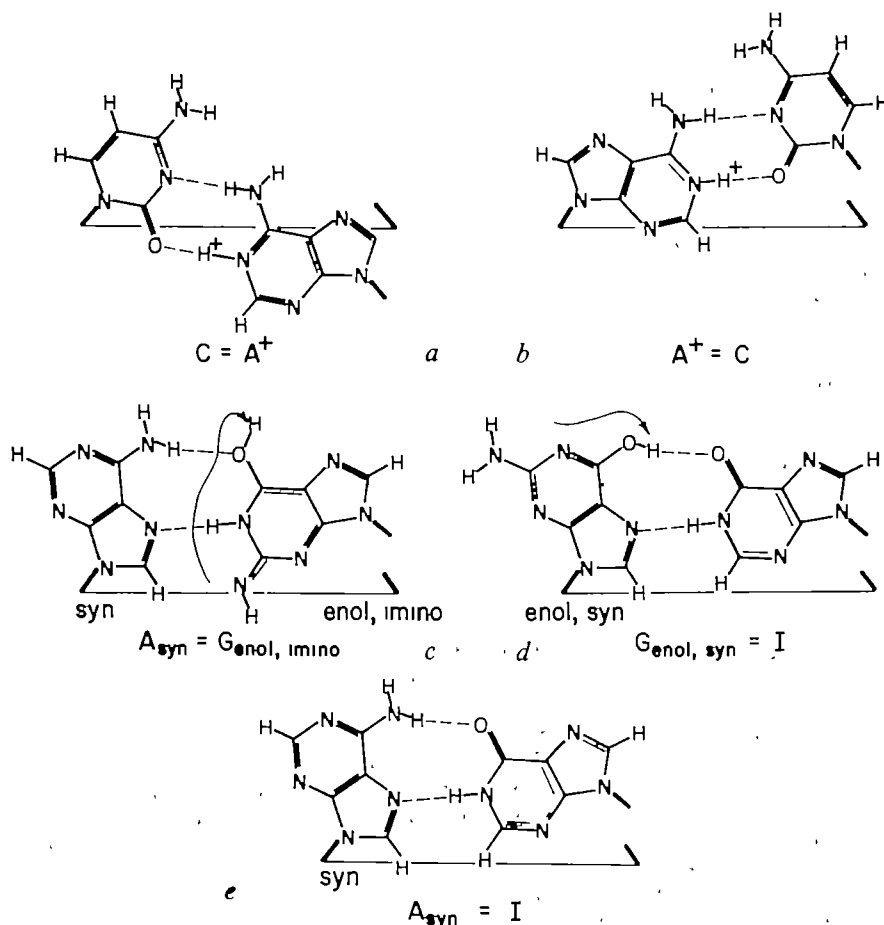
Fig. 3 Additional wobble-type base pairs for consideration in codon–anticodon interaction. A–C and C–A (*a,b*), which do not normally occur in codon–anticodon interaction, can only make this type of pair with two H bonds when one of its members is protonated. Among the pu–pu pairs, all involve one member in the syn configuration, but A_{syn}–I (*e*) is unique in not requiring an unfavourable tautomer. The requirement for an unfavoured tautomeric shift in the case of A_{syn}–$G_{enol, imino}$ (*c*) is not to enable the H bonding itself, but to remove an otherwise interfering interaction of the 2-amino of G with the C8 hydrogen of A.

Presumably, A–C and G–G do not serve at the wobble position because their pairing schemes require minor tautomers in high frequency, and G–A and C–A are irrelevant since A never occurs in the wobble position of tRNA[a].

U–U, C–C, U–C, C–U–. These do not occur in long Watson–Crick helices (J. Balcerski, M.D.T. and J.R.F., unpublished), and are forbidden by the code. Their occurrence would require a large backbone distortion that would enhance electrostatic repulsion of the phosphates brought closer together. The exception of U–U[a]–oxacetate pairing[aa] must be due to some unique effect of the substituent on C5 of the pyrimidine on the tRNA.

We conclude that schemes for pairing at the wobble position that require minor tautomers of bases[a,a] cannot account for

rise to complementary mispairs at all three codon positions. Mispairing at the wobble position could possibly also arise from the additional wobble pairs proposed in Fig. 3, since the ribosomal constraints must be relaxed somewhat to allow "normal" wobble at that position. Higher levels of miscoding due to mutations that reduce the ribosomal constraints or to drugs with similar consequences should be superimposed on this natural level of infidelity. Still higher levels of miscoding should occur *in vitro* when unnatural concentrations of cations or other artefactual conditions are used that overcome the specificity of the codon–anticodon interaction. Miscoding in this instance could involve an even wider range of non-Watson–Crick base oppositions, particularly if only two base pairs have to be stabilised.
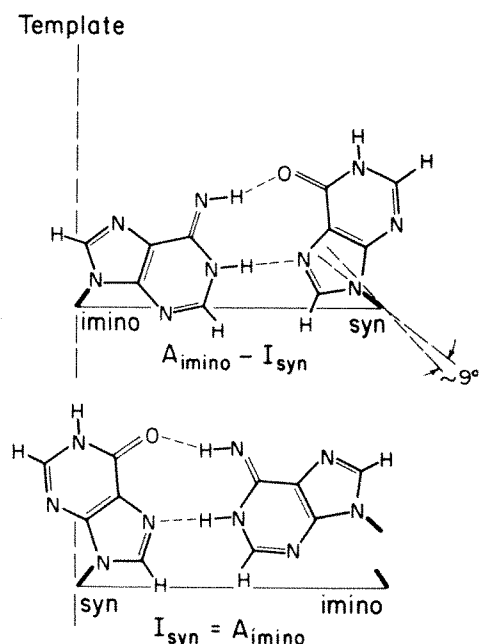
Template

Fig. 4 The steric relationship between a complementary pu–pu pair, when one residue has its locus fixed as part of a template, and either member is in the syn configuration. The particular example shown is that of an A–I pair with A as template (top) and I as template (bottom). Whereas the former meets complementary base-pair glycosyl bond separation distance without a significant dislocation of the backbone segment to which the non-template base is linked, wobble-type dislocation results when the syn residue is on the template.

Evaluation of the mispairing hypothesis

In vivo measurements of suppression of polar mutations in structural genes for β-galactosidase[30] and ornithine trans-carbamylase[8] and of the frequency of incorrect insertion of amino acids into rabbit haemoglobin and chick ovalbumin[31,32] lead to values of infidelity as high as 10^{-4} per amino acid site. This is consistent with what is expected from the frequency of the most favoured complementary mispairs, that is, those requiring only unfavoured tautomers (10^{-4}–10^{-5}). Thus, the *in vivo* infidelity is of a level that might be expected from the frequency of the minor tautomers in both message and tRNA before their interaction within the ribosome.

The *in vivo* level of infidelity is the sum of errors in several steps. Errors due to complementary mispairs will contribute a maximum frequency of 5×10^{-5} per residue in both transcription and codon–anticodon interaction. This is equivalent to an infidelity frequency of 3×10^{-4} per amino acid site, ignoring degeneracy of the code, which would lower this value. The indications are[33] that amino acid activation and aminoacylation together contribute another factor of 10^{-4}, for an overall error level of 4×10^{-4}, which is consistent with that measured *in vivo*. Although kinetic proofreading[34] of the codon–anticodon interaction may be necessary to screen out errors from such effects as two-base pair interactions, there is no obvious basis and no apparent need for proofreading of infidelity caused by complementary mispairs. This is because the mispairs are frozen within the ribosome and the level of infidelity to which they give rise is not greater than that contributed by the other steps of protein synthesis. Indeed, energy used for proofreading at this stage would seem wasteful.

If complementary mispairing is a source of errors in codon–anticodon interaction, there should be a defined pattern of miscoding which involves all complementary mispairs, that is, U–G, G–U, A–C, C–G and A–G at all codon positions, A–A, G–A and C–A at only the first and second positions (since

tRNAs never contain A in the wobble position), and G–I at only the third (since I only occurs at the wobble position). While pyr–pyr pairs should not occur at any codon position, additional A–C, A–G and G–I wobble-type pairs (Fig. 3) may possibly occur at the third position. The only *in vivo* coding data[30] confirms the occurrence of G–U, and A–C in both the second and third codon positions, with the frequency of background suppression due to A–C occurring at the 0.01 % level, which is consistent with the frequency of required tautomers. A higher value was observed for G–U, however; reasons for this are not apparent.

In vitro studies of codon–anticodon interaction do not provide a basis for evaluating the mispairing hypothesis. The errors occur at 10–100-fold or greater frequency than *in vivo*[24], presumably because of artefactual stabilisation of a variety of "nonspecific" interactions.

This investigation of base pairing in codon–anticodon interaction provides a basis for refinement of the wobble hypothesis and a chemical basis for infidelity in the process. Although our approach follows that which we used to investigate base pairing in nucleic acid biosynthesis[3], the present analysis is less satisfying because of the paucity of data against which to evaluate our hypothesis on base pairing in protein synthesis, but the concepts presented provide a basis for experimental focus.

This study has led us to consider the role of the ribosome in protein synthesis. It seems that the specificity of codon–anticodon interaction derives as much from the ribosomal constraints as from the limited degree of specificity that can come from the base-pairing interactions. The development of such constraints, which assure biologically acceptable levels of fidelity and efficiency of protein synthesis, may have been crucial during evolution of the ribosome.

1 Watson, J. D., and Crick, F. H. C., *Nature*, 171, 737–738 (1953).
2 Crick, F. H. C., *Proc. R. Soc.*, B167, 331–347 (1969).
3 Topal, M. D., and Fresco, J. R., *Nature*, 263, 285–289 (1976).
4 Crick, F. H. C., *J. molec Biol.*, 19, 548–555 (1966).
5 Fresco, J. R., and Topal, M. D., *Abstracts 100th Anniversary Meeting of Am. chem. Soc.*, New York, April 1976.
6 Staehelin, M., *Experientia*, 27, 1–11 (1971).
7 Lomant, A. J., and Fresco, J. R., *Prog. Nuc. Acid Res. molec. Biol.*, 15, 185–218 (1975).
8 Gorini, L., Jacoby, G. A., and Breckenridge, L., *Cold Spring Harb. Symp. quant. Biol.*, 31, 657–664 (1966).
9 Davies, J., *Cold Spring Harb. Symp. quant. Biol.*, 31, 665–670 (1966).
10 Gorini, L., in *Ribosomes* (edit. by Nomura, M., Tissieres, A. and Lengyel, P.), 791–803 (Cold Spring Harbor Labroatory, New York, 1974).
11 Tucker, G., and Irvin, J., *J. Am. chem. Soc.*, 73, 1923–1929 (1951).
12 Kenner, G. W., Reese, C. B., and Todd, A. R., *J. chem. Soc.*, 855–859 (1955).
13 Katritsky, A. R., and Waring, A. J., *J. chem. Soc.*, 1540–1544 (1962).
14 Wolfenden, R., *J. molec. Biol.*, 40, 307–310 (1969).
15 Jordan, F., and Pullman, B., *Theor. chim. Acta*, 9, 242–252 (1968).
16 Tinoco, I., Jr, Davis, R. C., and Jaskunas, S. R., in *Molecular Associations in Biology* (edit. by Pullman, B.), 77–92 (Academic, New York and London, 1968).
17 Jukes, T. H., *Nature*, 246, 22–26 (1973).
18 Barrell, B. G., and Clark, B. F. C., *Handbook of Nucleic Acid Sequences* (Joynson-Bruvvers, Oxford, 1974).
19 Gorini, L., and Beckwith, J. R., *A. Rev. Microbiol.*, 20, 401–422 (1966).
20 Sakore, T. D., and Sobell, H. M., *J. molec. Biol.*, 43, 77–87 (1969).
21 Davis, B. D., Anderson, P., and Sparling, P. F., *J. molec. Biol.*, 76,2 23–232. (1973)
22 Soll, D. *et. al.*, *Cold Spring Harb. Symp. quant. Biol.*, 31, 51–51 (1966).
23 Soll, D., Cherayil, J. D., and Bock, R. M., *J. molec. Biol.*, 29, 97–112 (1967).
24 Weinstein, I. B., Friedman, S. M. and Ochoa, M., Jr, *Cold Spring Harb. Symp. quant. Biol.*, 31, 671–681 (1966).
25 Yoshida, M., Takeishi, K., and Urita, T., *Biochem. biophys. Res. Commun.*, 39, 852–857 (1970).
26 Ohashi, Z., Saneyoshi, M., Harada, F., Hara, H., and Nishimura, S., *Biochem. biophys. Res. Commun.*, 40, 866–872 (1970).
27 Nirenberg, M., *et al.*, *Cold Spring Harb. Symp. quant. Biol.*, 31, 11–24 (1966).
28 Nishimura, S., *Prog. Nucl. Acid. Res. molec. Biol.*, 12, 49–85 (1972).
29 Bock, R. M., Soll, D., and Cherayil, J. D., *Ind. J. Biochem.*, 4 (Suppl.), 25 (1967).
30 Strigini, P., and Brickman, E., *J. molec. Biol.*, 75, 659–672 (1973).
31 Loftfield, R. B., and Vanderjagt, D., *Biochem. J.*, 128, 1353–1356 (1972).
32 Loftfield, R. B., *Biochem. J.*, 89, 82–92 (1963).
33 Hopfield, J. J., Yamane, T., Yue, V., and Coutts, S. M., *Proc. natn. Acad. Sci. U.S.A.*, 73, 1164–1168 (1976).
34 Hopfield, J. J., *Proc. natn. Acad. Sci. U.S.A.*, 71, 4135–4139 (1974).

Formation of stable crystalline enzyme–substrate intermediates at sub-zero temperatures

Anthony L. Fink & Ahmed I. Ahmed

Division of Natural Sciences, University of California, Santa Cruz, California 95064

Sub-zero temperatures can be used to trap intermediates in enzyme-catalysed reactions using suitable cryosolvents. The feasibility of obtaining such intermediates in the crystalline state for X-ray diffraction studies has been demonstrated with several proteases, using specific substrates and optimal pH.

Two problems in the study of enzyme mechanisms are the rapidity of the reaction and the low concentration of intermediates present in steady-state conditions. These cause technical difficulties in the determination of events occurring during the dynamic processes of catalysis. The existence and structure of intermediates in the catalytic reaction pathway must be known before the mechanism can be ascertained, and the great catalytic efficiency of enzymes can be explained. X-ray crystallography is one of few techniques that can yield structural data about such intermediates at the desired resolution. Unfortunately X-ray diffraction is limited in that it is essentially a static procedure and requires stable, crystalline compounds. Recent developments, however, suggest that this might be overcome at sub-zero temperatures.

It has been known for some time that many enzymes are catalytically active in aqueous organic solvent systems at sub-zero temperatures[1-6]. We have been exploiting different free energies and enthalpies of activation for the different steps in the overall enzyme-catalysed reaction to accumulate intermediates at very low temperatures. The basis of this approach is that when the reaction is initiated by mixing enzyme and substrate at a very low temperature the rates of intermediate transformations with sufficiently high activation energies will become negligible. Consequently, turnover will not occur and intermediates will accumulate. Thus instead of steady-state turnover, a series of reaction steps will be seen, each corresponding to the transformation of one intermediate into

another as the temperature increases. Eventually a temperature will be reached that is sufficiently high for turnover to occur. As long as the temperature is sufficiently low to avoid turnover, the enzyme will be trapped in the form of an intermediate or intermediates. The method is discussed in more detail elsewhere[9]. For several enzymes it has been possible to demonstrate the accumulation of intermediates found in the reaction with kinetically specific (good) substrates at optimal pH by this technique[3,6,10-15]. In these studies dissolved enzyme and various spectral techniques were used to detect and characterise the trapped intermediate.

We have investigated the feasibility of using sub-zero temperatures to obtain the trapped enzyme–substrate intermediates in crystalline form and sufficiently stable for high resolution X-ray diffraction studies. The intermediates used were acyl-enzymes formed from serine proteases. The reaction scheme for these enzymes can be represented by equation (1), in which ES is the initial non-covalent Michaelis complex, I_P and I_{NP}

$$E + S \rightleftharpoons ES \rightleftharpoons \left\{ \begin{array}{c} I_P \\ \updownarrow \\ I_{NP} \end{array} \right\} \overset{k_2}{\rightleftharpoons} EA \overset{k_3}{\longrightarrow} E + P_2 \qquad (1)$$
$$+$$
$$P_1$$
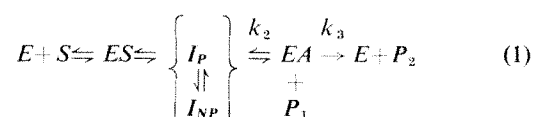
represent possible subsequent productive and non-productive intermediates respectively, EA is the acyl-enzyme intermediate and k_2 and k_3 are the rate constants for acylation and deacylation.

The serine proteases were chosen for several reasons. We have shown that sub-zero temperatures and 65% aqueous dimethyl sulphoxide or 70% aqueous methanol have no adverse effects on the structural and catalytic properties of chymotrypsin and trypsin[6,7]. We have also shown that essentially stoichiometric amounts of acyl-chymotrypsin, formed from specific p-nitrophenyl esters, can be isolated at sub-zero temperatures[6]. For specific ester substrates of chymotrypsin the acylation reactions can be orders of magnitude faster than

Table 1 Acylation and deacylation rates and acyl-enzyme concentrations for dissolved enzymes

Enzyme	Substrate	Solvent	pH*	Temperature (°C)	$[S_0]$ (M × 10³)	$[E_0]$§ (M × 10⁶)	EA† (%)	Acylation k_{obs} (s⁻¹)	Deacylation‡ v_i (mol s⁻¹)
Elastase	ZAP	70% MeOH	7.2	−43.4	3.2	13.4	82	1.1×10^{-3}	
Elastase	ZAP	70% MeOH	5.7	−50.1	0.75	5.1	92	3.2×10^{-4}	
Elastase	ZAP	70% MeOH	7.2	−37.0	1.5	5.6	79	1.8×10^{-3}	1.2×10^{-10}
Elastase	AcAla₃Me	70% MeOH	7.2	−52.7	1.9	1.6	47	2.3×10^{-4}	
α-Chymotrypsin	AcTrpP	60% DMSO	5.7	−41.5	1.0	6.6	86	5.0×10^{-4}	6.2×10^{-12}
δ-Chymotrypsin	AcTrpP	52% DMSO	5.7	−35.3	1.4	7.6	61	1.3×10^{-3}	4.6×10^{-11}
γ-Chymotrypsin	AcTrpP	52% DMSO	5.7	−41.2	1.5	17.0	44	7.6×10^{-4}	1.2×10^{-9}
γ-Chymotrypsin	AcTrpP	70% MeOH	7.0	−72.9	0.7	48.0	83	1.3×10^{-4}	9.1×10^{-10}
γ-Chymotrypsin	AcTrpP	70% MeOH	5.1	−43.6	4.1	6.1	82	6.6×10^{-4}	2.4×10^{-10}
γ-Chymotrypsin	AcTrpP	70% MeOH	6.7	−56.1	1.2	6.4	100	3.6×10^{-4}	9.6×10^{-11}
Trypsin	ZLysP	65% DMSO	7.7	−32.6	1.0	28.0	44	3.2×10^{-4}	4.5×10^{-10}

Substrate abbreviations: Z, N-carbobenzoxy; A, Ala, L-alanine; P, p-nitrophenyl ester; Ac, acetyl; Me, methyl; Trp, L-tryptophan; Lys, L-lysine. Solvents were 70% aqueous methanol (acetate buffer); 60 or 65% aqueous dimethyl sulphoxide (acetate) (DMSO); and 52% aqueous dimethyl sulphoxide, saturated with ammonium acetate at −80 °C (52% DMSO). Apparent protonic activity, pH*, of the aqueous organic solvents was measured at 25 °C and extrapolated to sub-zero temperatures using standard curves[41].
*Active-site normalities of the enzyme preparations were determined by standard procedures[31].
†Acyl-enzyme concentration expressed as fraction of total enzyme concentration. Estimated error ±15%.
§Observed initial rate of deacylation (turnover) (see text).

Table 2   Rates of acylation and deacylation, and active-site occupancies for crystalline enzymes

Enzyme	Substrate**	Crystal† size (µm)	Solvent**	pH*	Temperature (°C)	$[S_0]$ (M × 10³)	Solubility (M)	EA‡ (%)	Acylation k_{obs} (s⁻¹)	Deacylation k_{obs} (s⁻¹)
Elastase	ZAP	320	70% MeOH	7.2	−46.1	1.0	2.1×10^{-8}	86	2.7×10^{-5}	
Elastase	ZAP	1,200	70% MeOH	5.1	−47.8	0.6		≥55	7.8×10^{-6}§	0
Elastase	AcAla₃Me	80	70% MeOH	7.2	−54.5	2.8		≥86	1.35×10^{-5}	0
α-Chymotrypsin	AcTrpP	90	52% DMSO	6.1	−58.2	8.6	1.4×10^{-5}	3	$≥4.1 \times 10^{-5}$	0
δ-Chymotrypsin	AcTrpP	Amorphous	52% DMSO	6.1	−58.2	7.6	6.1×10^{-6}	82	$≥4.1 \times 10^{-5}$	0
γ-Chymotrypsin	AcTrpP	25	52% DMSO	5.8	−49.4	7.0	7.0×10^{-6}	≥61	$≥5.4 \times 10^{-5}$	0
Trypsin	ZLysP	70	52% DMSO	6.1	−58.2	9.8	2.0×10^{-5}	66	$≥4.1 \times 10^{-5}$	0

**See footnotes to Table 1.
†Longest dimension.
‡The reported values are minimum ones. Estimated errors are ±25%.
§This reaction appeared biphasic, the first part of the reaction with $k_{obs} = 2.4 \times 10^{-4}$ s⁻¹.

deacylation at temperatures below −40 °C in 65% aqueous dimethyl sulphoxide[6,7,10,11]. We have also detected and accumulated intermediates in addition to the acyl-enzyme in the reactions of chymotrypsin, elastase and trypsin with specific substrates[6,7,10,11]. Since our studies suggest that these intermediates are present, although not readily detected, when the reactions are carried out in "normal" conditions, knowledge of their structure is desirable[10].

Crystals of chymotrypsin and elastase have been shown to be catalytically active in aqueous solution at above-zero temperatures[16-23]. Preliminary studies on the effects of crystal size on the reaction rate and substrate diffusion in enzyme crystals have shown that catalytic rates and accessible active-site concentrations depend on crystal size[24-26]. Crystallographic studies have been carried out on acyl-enzymes formed by the reaction of nonspecific substrates with chymotrypsin[23], and on enzyme-'substrate' complexes formed by reaction of serine proteases with virtual substrates[22], peptide competitive inhibitors[17,18], peptide chloromethyl ketones[16] and protein inhibitors[27]. These investigations have greatly furthered the understanding of the mechanisms of action of these enzymes. But the fact that the derivatives used were sufficiently stable for X-ray diffraction implies that the relative positions of the enzymes' catalytic groups and the substrate were not optimal.

Petsko[28,29] has shown that X-ray diffraction studies of enzymes at sub-zero temperatures using aqueous–organic solvent systems have advantages, such as greatly decreased radiation damage. If X-ray diffraction techniques are to be applied to crystalline, trapped enzyme–substrate intermediates the following requirements must be met. At least 25 per cent of the enzyme must be in the form of the intermediate in order to attain sufficient intensity in electron density difference maps. Only free enzyme or the desired intermediate should be present. Mixtures of more than one type of intermediate are undesirable, and the crystalline intermediate, once formed, must be stable for a sufficiently long period for diffraction data to be collected, probably a minimum of 1–4 d for high resolution studies. We have shown, using sub-zero temperatures, that intermediates such as acyl- and glycosyl-enzymes satisfy the latter two requirements[6-8,10-13]. This and the accompanying article[30] confirm the validity of the cryoenzymological approach for obtaining detailed structural information on enzyme–substrate intermediates.

Acyl-A enzyme formulation in solution

p-Nitrophenyl ester substrates are well suited for studies of acyl-enzyme formation and breakdown in the serine proteases. The chromophoric nitrophenol moiety provides a simple means of monitoring the reaction. The smaller rate constant and larger energy of activation for the deacylation step mean that at appropriately low temperatures not only is the rate of acylation much faster than that of deacylation but also that the latter can be essentially zero.

The acylation and deacylation reactions of the following enzyme–substrate combinations were studied in the range −20 to −70 °C using aqueous dimethyl sulphoxide and aqueous methanol cryosolvents: α-, δ-, and γ-chymotrypsin with N-acetyl-L-tryptophan p-nitrophenyl ester (AcTrpP), trypsin with Nᵉ-carbobenzoxy-L-lysine p-nitrophenyl ester, and elastase with N-carbobenzoxy-L-alanine p-nitrophenyl ester (ZAP) and N-acetyl-L-alanyl-L-alanyl-L-alanine methyl ester.

Rates of acylation and deacylation (turnover) were determined for the p-nitrophenyl ester substrates by monitoring the release of p-nitrophenol at 340 or 350 nm. The concentration of acyl-enzyme formed was estimated from the amount of p-nitrophenol released in the first-order acylation reaction. For dissolved enzyme experiments, the enzyme stock solution was diluted 1:4 with cryosolvent at 0 °C and then added to the substrate solution at the desired sub-zero temperature.

For each enzyme, the rates of acyl-enzyme formation and breakdown, and the maximum concentration of acyl-enzyme formed were determined in various conditions. Some representative results of such experiments with dissolved enzymes are given in Table 1. In each case the acylation was first order and very much faster than deacylation. Deacylation was much too slow for the reactions to be followed to completion in a convenient time. In Table 1 the deacylation rates are reported in the form of the observed initial velocities (v_i). The substrate concentrations used were of the order of K_m; hence, dividing v_i by the acyl-enzyme concentration gives an approximate value of the microscopic deacylation rate constant, k_3 in equation (1).

For each enzyme there are conditions in which the rate of deacylation is sufficiently slow that negligible amounts of acyl-enzyme would be lost during the time necessary for the collection of diffraction data. For example, for elastase in 70% aqueous methanol at pH* 5.7 the amount of dissolved N-carbobenzoxy-L-alanyl-elastase which will hydrolyse in 24 h is 0.4% at −70 °C, 2.2% at −60 °C, and 11.0% at −50 °C. Acyl-enzyme concentrations approximately stoichiometric with the enzyme concentrations were only obtained at temperatures sufficiently low that no turnover was observed.

In conditions in which turnover occurs, the amount of enzyme converted into acyl-enzyme is predicted to be proportional to the substrate concentration unless it is much greater than K_m (ref. 31). This was observed to be the case. For example, with elastase and ZAP the fraction of total enzyme in the form of acyl-enzyme was a function of temperature; at increasingly higher temperatures, as the deacylation rate increased, the proportion of acyl-enzyme decreased. Consequently to obtain the maximum concentration of acyl-enzyme, it is essential to adjust pH* and temperature such that deacylation is negligible.

Experiments with dissolved α-, δ-, and γ-chymotrypsin indicate that approximately stoichiometric concentrations of acyl-enzyme can be obtained in the −50 °C to −70 °C range, using 65% aqueous dimethyl sulphoxide at pH* 5–8. Since both acylation and deacylation have similar pH dependencies, the relative ratio of rates of acylation to deacylation is essentially pH independent. In the conditions used, the rates of acylation were several orders of magnitude greater than that for deacylation. The rates of deacylation, especially at temperatures below −50 °C and at low pH*, were so slow that the acyl-enzymes could be made stable for a time scale of weeks or longer.

Because aqueous dimethyl sulphoxide cryosolvents are much

more viscous, they are less desirable than the corresponding methanol solvents[32]. The nucleophilicity of methanol towards acyl-chymotrypsins, however, is much greater than that of water[33], which suggested that in experiments at sub-zero temperatures deacylation would be much faster with aqueous methanol than with aqueous dimethyl sulphoxide for the same temperature. As Table 1 shows, with γ-chymotrypsin and AcTrpP the deacylation rate was about 15 times greater in the methanol solvent. This value is in good agreement with that calculated from the data of Bender *et al.*[33] obtained at 25 °C.

Crystalline acyl-enzyme formation

As preliminary attempts to crystallise the trapped acyl-enzymes at very low temperatures were unsuccessful, we started with the enzymes already in the crystalline state. Crystals of the enzymes were prepared as described before[34,35], except that γ-chymotrypsin was crystallised using 40 % saturated ammonium acetate, pH 5.6, rather than ammonium sulphate[34]. Crystals were grown in small capillary tubes, closed with dialysis membrane[36]. For most experiments with crystalline enzyme the procedure was as follows. One or more crystal(s) of the enzyme was added to the cryosolvent, either directly at 0 °C, or stepwise[29]. The crystal(s) was then added to the substrate solution at the desired temperature or vice versa. The reaction was either monitored directly, or samples were removed and analysed spectrophotometrically for *p*-nitrophenol. For experiments with elastase and the peptide ester substrate, acylation was followed by assaying for free elastase using ZAP at −40 °C. The solubilities of crystalline enzymes were determined by assaying samples from control experiments in which the substrate had been omitted. All experiments were carried out in conditions of excess substrate.

The precision of the values of the active-site occupancies of the crystals is not very high because of the difficulty of ascertaining the concentration of active sites in the crystal. The most accurate values were obtained when the crystalline acyl-enzyme deacylated at higher temperatures and was then dissolved in a suitable solvent for determination of total protein concentration. In conjunction with active-site normality titrations of the same batch of enzyme, this facilitated estimation of the active-site concentration in the crystal.

The data given in Table 2 are for experiments in which the enzymes were added in the crystalline state. With the exception of α-chymotrypsin, sufficient active-site occupancy was achieved for satisfactory X-ray diffraction studies. Crystalline and dissolved acyl-enzymes were shown to possess full potential catalytic activity when their temperatures were raised to 0 °C and the rates of turnover were measured.

Ammonium acetate was added to reduce the solubility of chymotrypsin and trypsin in 65 % aqueous dimethyl sulphoxide. Experiments with dissolved enzyme indicated that the catalytic reaction was essentially the same in both 65 % aqueous dimethyl sulphoxide (μ = 0.1 M) and in 52 % aqueous dimethyl sulphoxide, saturated (at −80 °C) with ammonium acetate, at the same temperature and pH*.

α-Chymotrypsin crystallises in a dimeric form in which the active site is blocked by the adjacent molecule. This packing geometry hinders access by any but the smallest of substrates or inhibitors[16,38]. Our observation that only 3 % of the active sites of crystalline α-chymotrypsin was occupied using AcTrpP compared with 20 times as much with γ-chymotrypsin in similar conditions, not only is in accord with expectations but serves as a valuable control. Active-site occupancies with crystalline α-chymotrypsin of 15–20 % at 25 °C have been claimed with similar substrates[19]. Attempts to crystallise δ-chymotrypsin have been unsuccessful[37], and so it was used in the amorphous state. The high active-site occupancy observed with this physical form of the enzyme is in accord with predictions that in very small crystals diffusion into the crystal is not limiting[24-26]. The active-site occupancy obtained with γ-chymotrypsin crystals is of the desired order for X-ray diffraction experiments.

As might be expected the results with trypsin are very similar to those for γ-chymotrypsin. For the crystalline enzyme the active-site occupancy and acylation–deacylation rate ratio indicate the suitability of trypsin for diffraction studies using aqueous dimethyl sulphoxide and temperatures below −50 °C. The explanation for the lower active-site occupancies with γ-chymotrypsin and tryspin is probably due to errors in determining the concentration of the active enzyme in the crystals.

For elastase a polypeptide ester substrate as well as a *p*-nitrophenyl ester was used. Both the molecular size and the kinetic parameters, k_{cat} and K_m, are very similar for ZAP and *N*-Ac-Ala-Ala-Ala methyl ester[39,40]. We have also found that non-covalent enzyme–substrate intermediates involving elastase and *N*-succinyl-Ala-Ala-Ala *p*-nitroanilide can be accumulated in a similar fashion to the acyl-enzyme (unpublished results). Most of the elastase experiments were carried out using aqueous methanol. The solubility of elastase was found to be much less than that of trypsin and chymotrypsin in the cryosolvents used. The results from experiments using crystalline elastase indicated that the active site occupancy for the peptide substrate was at least equal to or better than that for the *p*-nitrophenyl ester. The acylation reaction for the latter substrate with the largest enzyme crystals seemed biphasic (Table 2). The data obtained show that substantial concentrations of crystalline acyl-elastases can be obtained in 70 % aqueous methanol at the pH-optimum and that at temperatures below −60 °C the acyl-enzymes will be stable for several days. Good agreement in the observed rates of acylation and deacylation for crystalline elastase with ZAP was found in the results reported here and those obtained in the diffraction experiments[30]

Previous studies have indicated that at 25 °C the rates of reactions catalysed by crystalline enzymes range from 100 to 0.3 % of the rate with the enzyme in solution[24-26]. At sub-zero temperatures we observed substantially slower acylation when the crystalline enzymes were used than in the corresponding reaction with dissolved enzyme. For example, acylation was approximately 20 times slower with the crystalline enzyme for elastase and about 15 times slower for γ-chymotrypsin. As Table 2 shows, no deacylation was observed with the crystalline acyl-enzymes over 2–4-d periods at temperatures below −40 °C. This observation is in contradiction to the predicted deacylation rates based on the experiments using the dissolved acyl-enzymes (Table 1) and suggests that the intrinsic rate of deacylation is slowed down in the crystal by an amount at least comparable with that of the acylation. One interpretation of this result —that the crystal lattice forces are responsible and perhaps hinder the deacylation by impeding a conformational change— is supported by the observation that in crystalline trimethyl-acetyl-chymotrypsin (ambient temperature, aqueous buffer) the rate of deacylation is 50–100 times slower than in solution, and that the crystals shatter after 5–10 % deacylation (F. J. Kézdy, personal communication). These observations are not in accord with the report that indoleacryloyl-chymotrypsin had identical rates of hydrolysis when in solution or crystalline phase[42].

In conclusion, the results of this investigation show that temperatures in the −50 °C to −70 °C range (depending on the pH*), in conjunction with suitable cryosolvents, can be used to accumulate essentially stoichiometric concentrations of dissolved acyl-enzymes from serine proteases and specific substrates. Furthermore, substantial concentrations of acyl-enzymes in the crystalline state can be obtained in a similar manner. These experiments can be done at pH values corresponding to maximum activity. The accumulated acyl-enzymes (dissolved or crystalline) are stable for periods of days to months. By the choice of appropriate pH* and temperature the active-site occupancies and negligible deacylation rates necessary for X-ray diffraction studies can be attained. Confirmation of the results and predictions of this study are given in the accompanying article[30] concerning crystallographic studies of an acyl-elastase.

Alber for carrying out preliminary experiments and for stimulating discussions, and the NSF and NIH for financial support.

1 Maier, V. P., Tappel, A. L., and Volman, D. H., *J. Am. chem. Soc.*, 77, 1278–1280 (1955).
2 Bielski, B., and Freed, S., *Biochim. biophys. Acta*, 89, 314–323 (1964).
3 Douzou, P., Sireix, R., and Travers, F., *Proc. natn. Acad. Sci. U.S.A.*, 66, 787–792 (1970).
4 Douzou, P., and Letterier, F., *Biochim. biophys. Acta*, 220, 338–340 (1970).
5 Douzou, P., *Biochimie*, 53, 1135–1145 (1971); *Molec. cell. Biochem.*, 1, 15–27 (1973).
6 Fink, A. L., *Biochemistry*, 12, 1736–1742 (1973); *Archs Biochem. Biophys.*, 155, 473–474 (1973).
7 Fink, A. L., *Biochemistry*, 13, 277–279 (1974); *J. biol. Chem.*, 249, 5027–5032 (1974).
8 Fink, A. L., and Wildi, E., *J. biol. Chem.*, 249, 6087–6089 (1974).
9 Fink, A. L., *J. theor. Biol.* (in the press).
10 Fink, A. L., *Biochemistry*, 15, 1580–1586 (1976.)
11 Fink, A. L., McGarraugh, G., and Farzami, B., *J. biol. Chem.* (in the press).
12 Fink, A. L., and Good, N., *Biochem. biophys. Res. Commun.*, 58, 126–131 (1974).
13 Fink, A. L., and Angelides, K. J., *Biochem. biophys. Res. Commun.*, 64, 701–708 (1975).
14 Balny, C., and Hastings, J. W., *Biochemistry*, 14, 4719–4723 (1975).
15 Hastings, J. W., Balny, C., LePeuch, C., and Douzou, P., *Proc. natn. Acad. Sci. U.S.A.*, 79, 3468–3472 (1973).
16 Segal, D. M., Powers, J. C., Cohen, G. H., Davies, D. R., and Wilcox, P. E., *Biochemistry*, 10, 3728–3738 (1971).
17 Robertus, J. D., Kraut, J., Alden, R. A., and Birktoft, J. J., *Biochemistry*, 11, 4293–4303 (1972).

18 Shotton, D. M., White, N. J., and Watson, H. C., *Cold Spring Harb. Symp. quant. Biol.*, 36, 91–105 (1971).
19 Kallos, J., *Biochim. biophys. Acta*, 89, 364–366 (1964).
20 Sigler, P. B., and Skinner, H. C. W., *Biochem. biophys. Res. Commun.*, 13, 236–240 (1963).
21 Sigler, P. B., Jeffery, B. A., Matthews, B. W., and Blow, D. M., *J. molec. Biol.*, 15, 175–192 (1966).
22 Steitz, T. A., Henderson, R., and Blow, D. M., *J. molec. Biol.*, 46, 337–348 (1969).
23 Henderson, R., *J. molec. Biol.*, 54, 341–354 (1970).
24 Doscher, M. S., and Richards, F. M., *J. biol. Chem.*, 238, 2399–2406 (1963).
25 Quiocho, F. A., and Richards, F. M., *Biochemistry*, 5, 4062–4076 (1966).
26 Sluyterman, L. A. Æ., and de Graaf, M. J. M., *Biochim. biophys. Acta*, 171, 277–287 (1969).
27 Blow, D. M., *et al.*, *J. molec. Biol.*, 69, 137–144 (1972).
28 Marsh, D. J., and Petsko, G. A., *J. appl. Crystallogr.*, 6, 76–80 (1973).
29 Petsko, G. A., *J. molec. Biol.*, 96, 381–392 (1975).
30 Alber, T., Petsko, G. A., and Tsernoglou, D., *Nature*, 263, 297–300 (1976).
31 Bender, M. L., *et al.*, *J. Am. chem. Soc.*, 88, 589–598 (1966).
32 Travers, F., Douzou, P., Pederson, T., and Gunsalus, I. C., *Biochimie*, 57, 43–48 (1975).
33 Bender, M. L., Clement, G. E., Gunter, C. R., and Kézdy, F. J., *J. Am. chem. Soc.*, 86, 3697–3703 (1964).
34 Kunitz, M., *J. gen. Physiol.*, 22, 207–237 (1938).
35 Shotton, D. M., *et al.*, *J. molec. Biol.*, 32, 155–156 (1968).
36 Zeppezauer, M., Eklund, H., and Zeppezauer, E. S., *Archs biochem. Biophys.*, 126, 563–573 (1968).
37 Kraut, J., Wright, T. H., Kellerman, M., and Freer, S. T., *Proc. natn. Acad. Sci. U.S.A.*, 58, 304–311 (1967).
38 Fersht, A. R., and Renard, M., *Biochemistry*, 13, 1416–1426 (1974).
39 Geneste, P., and Bender, M. L., *Proc. natn. Acad. Sci. U.S.A.*, 64, 683–685 (1970).
40 Gertler, A., and Hoffmann, T., *Can. J. Biochem.*, 48, 384–386 (1970).
41 Maurel, P., Hui Bon Hoa, G., and Douzou, P., *J. biol. Chem.*, 250, 1376–1382 (1975).
42 Rossi, G. L., and Bernhard, S. A., *J. molec. Biol.*, 49, 85–91 (1970).

Crystal structure of elastase–substrate complex at −55 °C

T. Alber*, G. A. Petsko & D. Tsernoglou

Department of Biochemistry, Wayne State University School of Medicine, Gordon H. Scott Hall of Basic Medical Sciences, 540 East Canfield Avenue, Detroit, Michigan 48201

The structure of a specific acyl-enzyme intermediate in the elastase-catalysed hydrolysis of N-carbobenzoxy-L-alanyl-p-nitrophenol ester has been determined by X-ray diffraction at 3.5 Å resolution. The acyl-enzyme was stabilised by cooling the crystal to −55 °C during substrate addition and data collection.

THE structure of a specific acyl-enzyme intermediate in the catalytic reaction of a serine protease has been determined by the X-ray diffraction at sub-zero temperatures. The use of a fluid cryoprotective solvent methanol–water, allowed addition of the substrate to the crystalline enzyme at −55 °C. At this temperature an acyl-enzyme formed within 1 d but was stable for more than a week. This work demonstrates the potential of low temperature protein crystallography for directly determining the details of the interactions of enzymes with their actual substrates.

Cryoenzymology

Protein crystallography is the most powerful technique for observing the interactions of enzymes with small molecules at atomic resolution. Since protein crystals contain intermolecular channels filled with the liquid of crystallisation, it is possible to diffuse small molecules into the crystals. There they can bind to the protein molecules if their binding sites are not blocked by molecule–molecule contacts. The crystal structure of the complex can then be determined by the usual techniques. The interactions of inhibitors and pseudo-substrates with crystalline enzymes have been

studied in this way[1,2]. Until now, however, it has not been possible to use this approach to study the binding of an actual substrate, because collection of a full set of three-dimensional X-ray diffraction data at high resolution takes at least several days, while the lifetime of a productive enzyme–substrate complex, even in the crystalline state, is several seconds or less. Thus, averaged over the time of data collection, the observed structure will contain only a negligible amount of enzyme–substrate complex.

The structure of a productive enzyme–substrate complex or an intermediate on the catalytic pathway could be determined crystallographically if the rate of enzyme catalysis could be reduced considerably. Douzou has pointed out that this can be done by lowering the temperature[3]. The Arrhenius equation relating reaction rates to temperature predicts that an enzymatic reaction having an energy of activation of −15 kcalorie mol^{-1} will be slowed by a factor of 10^8 on cooling from room temperature to −100 °C. At this low temperature an enzyme–substrate complex with a lifetime of 10^{-3} s at +25 °C would be stable for 1.2 d. Douzou and his colleagues have made use of this effect to study some enzyme–substrate interactions in solution spectroscopically[4] and have also determined the physicochemical properties of many aqueous organic mixed solvent systems[5]. These solvents remain fluid at very low temperatures and, if their physicochemical parameters are properly adjusted, do not denature enzymes nor alter the rate-determining reaction step[6,7].

Recently Fink and his coworkers made a detailed study of the reactions of several serine proteases in both solution and crystalline states with ester and amide substrates in aqueous organic solvents at sub-zero temperatures[10,11]. They concluded that in carefully controlled conditions the enzymes would function normally in this bizarre environment but
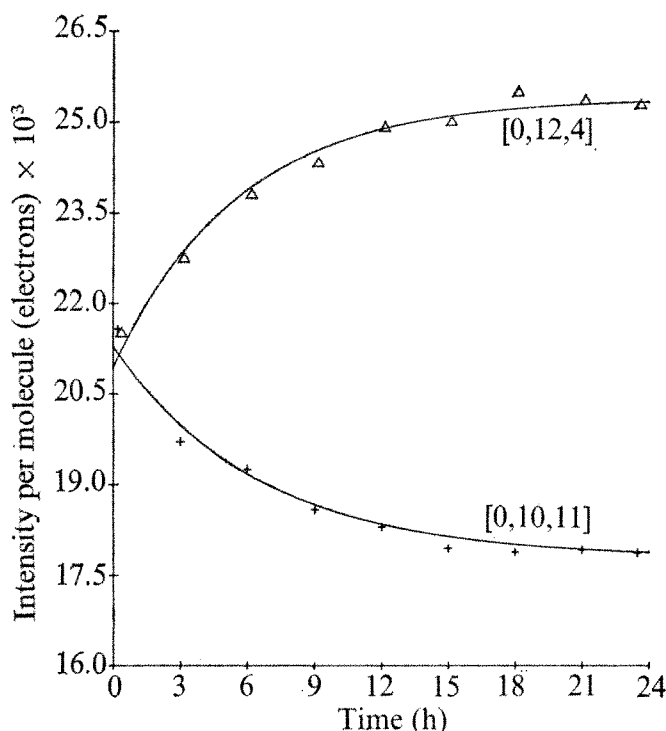
Fig. 1 Plots of the changes in intensity with time of two high-resolution 0*kl* reflections of elastase as substrate is added. The solid curves are calculated assuming a first-order rate constant of $5.4 \times 10^{-5}\,\text{s}^{-1}$.

their rates were so reduced as to make direct observation of short lived proteolysis intermediates feasible. Fink has also determined conditions in which sufficiently high concentrations of such intermediates can be accumulated in crystals and has characterised the species formed[11]. These studies are invaluable for planning and interpreting X-ray diffraction experiments at low temperatures.

We have extended the techniques of cryoenzymology to protein crystallography by cooling protein crystals immersed in fluid mixed solvents, and then flowing in substrate at low temperature. In this paper we report the results of such an experiment with a specific synthetic substrate of porcine pancreatic elastase at −55 °C in 70% methanol–30% water *p*H* 5.2 (*p*H* denotes protonic activity in mixed solvents at the specified temperatures[5,11,13]).

Cryoprotection of elastase crystals

We chose a serine protease for our first experiments because there is a large body of kinetic and structural data available for this class of enzymes, but many details of their mode of action are still disputed[8]. They are also attractive because the rate-determining step in the reaction differs for different types of substrate. This offers the possibility of trapping several different intermediates crystallographically at low temperatures[10]. We chose elastase as the particular crystalline enzyme to study because of its low solubility and high stability in alcohol–water mixtures[11] and the very favourable molecular packing in its crystals. The active site of the enzyme faces a solvent-filled interstice 18 Å in diameter, allowing the binding of large peptide substrates[12].

Elastase crystals are grown from sodium sulphate–sodium acetate solutions at low ionic strength. They are normally stabilised in 1.2 M sodium sulphate, 0.01 M acetate (*p*H 5.0) and it was in this salt-rich mother liquor that the room temperature crystal structure was solved by Watson and his colleagues[12]. To cool the elastase crystals to temperatures in the −50 °C range (which we knew to be sufficiently low to stabilise an enzyme–substrate complex from the work of Fink and his colleagues[11]), it was first necessary to find a

replacement mother liquor which would not freeze[13].

We found that elastase crystals are stable in a wide variety of mixed solvents, provided certain procedures are followed. The crystals cannot be transferred directly from 1.2 M sodium sulphate to alcohol–water mixtures, because of the low solubility of sodium sulphate in alcohol. The first step of the transfer is therefore to place the crystals in 0.01 M sodium acetate (*p*H 5.0). In this salt-free liquid they are stable for weeks at 4 °C. Transfer cannot be made directly to solvents of high alcohol concentration, as this cracks and partially disorders the crystals. Transfer can safely be made, however, to solvents of low alcohol concentration, and the percentage of alcohol may be increased to any desired level provided the gradual addition of alcohol is coupled with a gradual reduction in temperature[13]. Alcohol concentration and temperature are varied so as to keep the dielectric constant of the solvent as near as possible to the value for pure water[4,13]. It is important at each step to vary the *p*H of the aqueous buffered component to keep the protonic activity of the solvent constant[4,13]. With this procedure alcohol concentrations from 10 to 80% can be reached with no damage to the crystals. Since a number of different solvents were available to us, we based our decision on the normality of kinetic behaviour of the enzyme in the solvent and its fluidity at low temperatures. Our choice was methanol–water (70%: 30%, v/v), a mixture of very low freezing point with a viscosity very close to pure water, even at −75 °C (ref. 5).

Data collection and substrate binding

To obtain stable enzyme–substrate complexes, the substrate must be flowed into the crystal at sub-zero temperature. Richards and his colleagues haved described a flow cell for use with single-crystal X-ray diffractometers that holds the crystal firmly in place while allowing liquid to flow through it continuously[14]. We have used this device both for substrate binding and solvent transfer experiments. In practice, the crystal is mounted in the flow cell in 0.01 M acetate (*p*H 5.0). It is then cooled to 5 °C on the diffractometer and the liquid is replaced by 20% methanol–80% water at the same protonic activity. The temperature is then lowered 10 °C and the liquid is replaced with one of 10% higher methanol concentration. It takes about 7 min for the wave front of fresh liquid to reach the crystal, and equilibration at each temperature/alcohol stage takes 10 to 15 min. For substrate binding at −55 °C a flow rate of about 2 ml d⁻¹ was used, with the substrate dissolved in 70% methanol and flowed over the crystal in the usual way.

Data were collected on a Syntex P2₁ diffractometer specially modified for protein crystallography (G.A.P. and D.T., unpublished), equipped with a Syntex LT-1 low temperature device which was also extensively modified (G.A.P., unpublished). Data collection was by Wyckoff step scan[14], which enabled us to collect 2,400 reflections per day. All low temperature work was done at −55° ±4 °C.

The substrate chosen for this study was *N*-carbobenzoxy-L-alanyl-*p*-nitrophenyl ester (ZAP), as this was the substrate about which the most kinetic information was available[11]. It was dissolved in 70% methanol, 0.01 M acetate (*p*H* 5.2) at 3×10^{-3} M. To avoid problems due to its spontaneous hydrolysis the solution was made fresh twice daily and kept at −35 °C.

The use of the diffractometer–flow cell system made it possible to follow the binding of ZAP crystallographically. A small set of medium intensity reflections at high resolution were monitored at regular intervals. Plots of their intensity with time were monotonic and could be interpreted in terms of a first-order process with an overall rate constant of $5.4 \pm 1.3 \times 10^{-5}\,\text{s}^{-1}$ (Fig. 1). This is in excellent agreement with the acylation rate constant measured for the ZAP–elastase crystal system at −50 °C by Ahmed and Fink[11], if the value is adjusted for the difference in substrate

Table 1 Refined heavy-atom parameters and phasing statistics

Derivative	Coordinates*			Occupancy†	Temperature factor A^2	Rc‡	$\langle f_H \rangle$§	E‖
	x	y	z					
Uranyl nitrate	0.1170	0.5851	0.3962	92	33.1	0.48	88	40
(5 mM, 2 weeks soak at +25 °C)								
Sodium mersalyl	0.0311	0.2666	0.0637	65	25.8	0.53	104	62
(3 mM, 2 weeks soak at +25 °C)	0.0449	0.9713	0.2125	30	10.8			
	0.1763	0.7747	0.0082	61	155.4			

*x, y and z are the heavy-atom coordinates in one asymmetric unit (one elastase molecule) in fractions of the unit cell edges.
†Occupancy in electrons on an approximate absolute scale.

‡Rc is the Cullis R-factor $\dfrac{\Sigma||(F_{PH}-F_P)|-f_H|}{\Sigma|(F_{PH}-F_P)|}$ the sums being over all centric reflections[19].

§$\langle f_H \rangle$ is the r.m.s. heavy-atom structure amplitude in electrons.
‖E is the r.m.s. lack of closure of the phase triangle at the most probable phase, in electrons[20].
The mean figure-of-merit[21] for the 2,981 reflections to 3.5 Å is 0.73.

concentration (we observed that the rate varied linearly with substrate concentration, as expected). Collection of a complete set of three-dimensional data to 3.5-Å resolution was begun after the binding curves had levelled off; two days were required to collect the set. Neither low temperature nor substrate binding caused any damage to the elastase crystals; indeed, the mosaic spread of the crystal decreasing by 25% on cooling. The unit cell parameters of the enzyme in alcohol at −55 °C and the enzyme–substrate complex at the same conditions agreed within 0.5% with those of the enzyme in water at room temperature ($a=51.5$ Å, $b=58.0$ Å, $c=75.5$ Å, $\alpha=\beta=\gamma=90°$, space group $P2_12_12$), suggesting that the three cases were mutually isomorphous.

Phase determination

The crystallographic binding curves suggested that substrate was binding to the crystalline enzyme at low temperature, but to interpret the binding, a set of phase angles for the native enzyme structure was required. Phases were available from the earlier room temperature structure determination[12], but we did not use them. The room temperature structure was of an inhibited enzyme; moreover, one of the heavy-atom derivatives used to solve the structure was itself a covalent inhibitor. The phase angles from this structure determination therefore might be biased towards the presence of material in the substrate binding site. We decided to obtain phase information free from possible
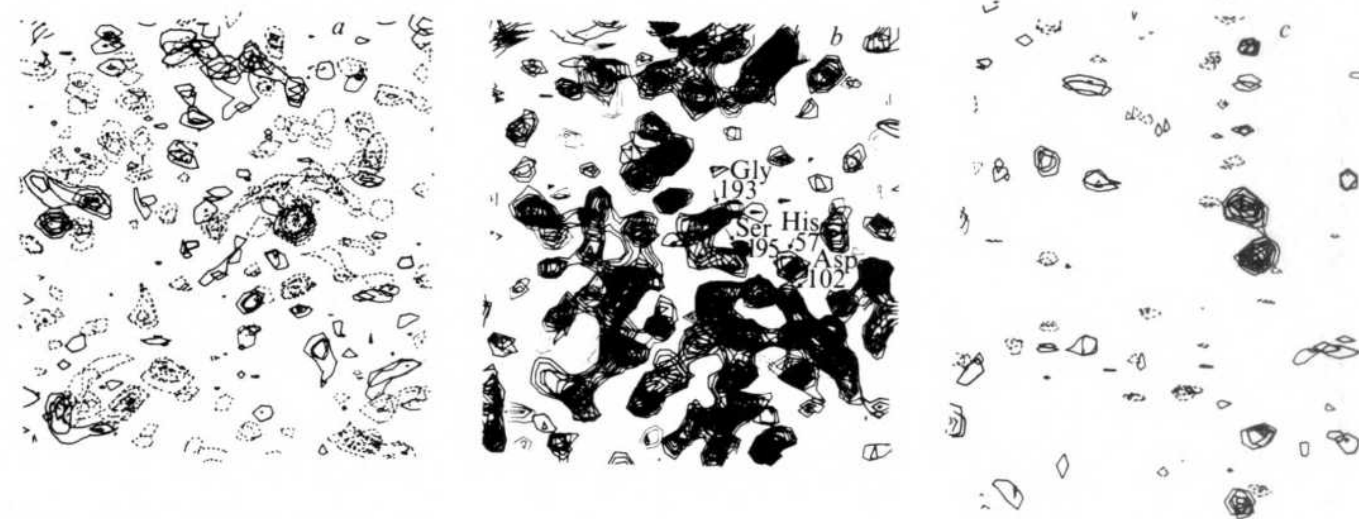
Fig. 2 Photographs of the active site region ($y = 10/60$ to $50/60$, $z = 10/60$ to $50/60$) of stacks of seven sections ($x = −13/60$ to $−7/60$) of the various electron density maps of elastase at low temperature. Phase angles derived for crystals at room temperature in sodium sulphate were used for all maps. For the difference maps the r.m.s. error in electron density was calculated from the revised Henderson and Moffat formula[18], and contours were drawn at intervals of this value, beginning at twice the error. Positive density is enclosed by solid contours, negative density by dashed lines. The ordinary electron density maps have been contoured at equal but arbitrary intervals of positive density; the interval was chosen to approximate that used by Watson *et al.* in Fig. 5 of ref. 12. *a*, Difference electron density map using as amplitudes F (elastase in methanol at −55 °C) −F (elastase in sodium sulphate at +25 °C). The spherical negative density is due to the replacement of a bound sulphate ion by acetate (T.A. and G.A.P., unpublished). The molecular boundary can be determined by reference to Fig. 2*b*. *b*, The amplitudes used in this map were $2F$ (elastase at −55 °C) −F (elastase at +25 °C). This gives a picture of the structure of the enzyme in methanol at −55 °C. Some of the important side chains in the active site are labelled. The black dot marks the position of the β-O of SER-195. (Compare with Fig. 5 of ref. 12.) *c*, Difference electron density map calculated with amplitudes F (elastase + ZAP at −55 °C) −F (elastase at −55 °C). The use of elastase in methanol at −55 °C as the parent structure in this difference map cancels out the high noise in the solvent region seen in Fig. 2*a*. *d*, The difference map of *c* superimposed on *b*, showing the interactions the substrate makes with the enzyme. The methyl group of the alanyl portion of the substrate points into the paper, where it is in contact with the methyl group of VAL-216.

bias by solving the structure of the uninhibited enzyme at 3.5-Å resolution using heavy-atom derivatives which do not bind in the substrate binding site. Two were quickly found, one of which, uranyl nitrate, was used in the earlier study[12]. All phase determination was done on the enzyme in 1.2 M sodium sulphate (pH 5.0) at +25 °C for reasons of speed and convenience. Details of the phasing will be published separately. Though not as high in quality as the earlier work, the phase determination was acceptable (Table 1). These phases for the uninhibited native enzyme in sodium sulphate were used for all electron density calculations.

Electron density maps

In addition to the room temperature data collected to phase the structure amplitudes, complete sets of data to 3.5-Å resolution were collected on elastase in 70% methanol at −55 °C and elastase + ZAP in 70% methanol at −55 °C. An electron density map of the native enzyme at +25 °C in sodium sulphate was virtually identical to that obtained by Watson and colleagues on the inhibited enzyme[12]. The only significant differences were the absence of the inhibitor density, the presence of a dense spherical feature in the active site interpreted as a bound sulphate ion[15], and the repositioning of HIS-57 to form a complete charge-relay system identical to that observed in chymotrypsin[16] (the histidine is prevented from assuming this position when the inhibitor is bound).

Two electron density maps were calculated for native elastase in methanol at −55 °C. The first, a difference electron density map between the low temperature and room temperature crystals, has as its largest feature a hole at the position of the bound sulphate ion, indicating that this ion is lost in the salt-free mixed solvent, as expected (Fig. 2a). There are no significant features within the molecular envelope, although there are numerous peaks and holes in the solvent-filled interstice. The only significant positive features are on the surface of the molecule, and we have interpreted these in some cases as slight shifts in the position of side chains, in others as the appearance of side-chain density where none was visible at room temperature (presumably because of thermal vibration). The second map shows the electron density of elastase in 70% methanol at −55 °C. The positions of important amino acids in the active site region were noted in this map (Fig. 2b).

We next calculated a difference Fourier synthesis which showed the difference in electron density between elastase and elastase + ZAP, both in 70% methanol at −55 °C. The only significant feature was a bilobed peak in the active site region (Fig. 2c). One of the lobes was within covalent bond distance from the β-oxygen of SER-195; the other, a planar density feature, was interpreted as the benzene ring of the carbobenzoxy group. There was no density for the p-nitrophenol portion of the substrate. An integrated electron count of the peak indicated approximately 80% occupancy for the substrate, in excellent agreement with the value found by Fink and Ahmed[11]. When this difference map was superimposed on the electron density map of the unsubstituted enzyme at −55 °C, the close contact with SER-195 was obvious (Fig. 2d; the black dot in the centre of the photograph is the β-oxygen position as determined by Watson and his colleagues at 2.5-Å resolution[17]). We interpret these maps as depicting a stable, productive acyl-enzyme intermediate in the catalytic process, as predicted by Fink and Ahmed in the accompanying article[11]. We believe that the substrate was bound to the crystalline enzyme, that p-nitrophenol was split off when the alanyl residue became covalently attached to the β-oxygen of SER-195, and that deacylation was prevented from occurring by the low temperature used in the study. To test this interpretation, we attempted to wash the crystal free of bound substrate at −55 °C. After 4 d of continuous flow of substrate-free 70% methanol a set of data was collected

and another difference Fourier synthesis was calculated. The results showed no reduction in the amount of bound substrate, suggesting that the mode of attachment was covalent. When the temperature was raised to −10 °C, the intensity binding curves reversed, suggesting loss of bound substrate. Another set of data was collected after 20 h at the elevated temperature, and this showed no detectable substrate remaining in the active site.

Low temperature protein crystallography

We do not believe it is appropriate, at this stage of the analysis with its modest resolution, to use these results to answer detailed questions about the mechanism of serine protease catalysis or the nature of the catalytic power of enzymes. We hope that some insights into these questions will come from the extension of this work to much higher resolution and several different substrates. For example, since the rate-limiting step for the hydrolysis of p-nitro-anilide substrates is the formation of the acyl enzyme rather than its breakdown, binding a substrate of this type to the crystalline enzyme at low temperature may yield a pre-acyl enzyme intermediate or the Michaelis complex[11]. This would give a second "stop-action" photograph of a different point in the overall reaction pathway. Our present study, combined with the work of Fink and Ahmed in the accompanying article[11], which was essential for the planning and interpretation of our experiments, demonstrates that it is feasible to use low temperature protein crystallography to stabilise and study actual enzyme–substrate complexes. We believe that the technique has great potential to reveal the atomic details of enzyme catalysis, especially when it can be applied to several substrates with different rate-determining steps. In conjunction with kinetic and spectro-scopic studies of crystalline and dissolved enzymes[22,26], this technique offers the hope of increased understanding of the events in the catalysis of reactions by enzymes. We believe that cryoenzymology can be applied to a wide variety of crystalliseable enzymes.

1 Phillips, D. C., *The Harvey Lectures, Series 66* (Academic, London, 1972).
2 Blow, D. M., and Steitz, T. A., *A. Rev. Biochem.*, **39**, 63–100 (1970).
3 Douzou, P., Sireix, R., and Travers, F., *Proc. natn. Acad. Sci. U.S.A.*, **66**, 787–792 (1970).
4 Douzou, P., *Biochimie*, **53**, 17–23 (1971).
5 Douzou, P., *Biochimie*, **53**, 1135–1145 (1971).
6 Douzou, P., *Molec. cell. Biochem.*, **1**, 15–27 (1973).
7 Douzou, P., Hui Bon Hoa, G., and Petsko, G. A., *J. molec. Biol.*, **96**, 367–380 (1975).
8 Hunkapiller, M. W., Smallcombe, S. H., Whitaker, D. R., and Richards, J. H., *Biochemistry*, **12**, 4732–4742 (1973).
9 Robillard, G., and Shulman, R. G., *J. molec. Biol.*, **86**, 541–558 (1974).
10 Fink, A. L., *Biochemistry*, **15**, 1580–1586 (1976).
11 Fink, A. L., and Ahmed, A. I., *Nature*, **263**, 294–297 (1976).
12 Watson, H. C., Shotton, D. M., Cox, J. M., and Muirhead, H., *Nature*, **225**, 806–811 (1970).
13 Petsko, G. A., *J. molec. Biol.*, **96**, 381–392 (1975).
14 Wyckoff, H. W., *et al.*, *J. molec. Biol.*, **27**, 563–578 (1967).
15 Shotton, D. M., White, N. J., and Watson, H. C., *Cold Spring Harb. Symp. quant. Biol.*, **36**, 91–105 (1971).
16 Blow, D. M., Birktoft, J. J., and Hartley, B. S., *Nature*, **221**, 337–340 (1969).
17 Sawyer, L., Shotton, D. M., and Watson, H. C., *Biochem. biophys. Res. Commun.*, **53**, 944–951 (1973).
18 Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C., and Tjian, R., *J. molec. Biol.*, **88**, 349–371 (1974).
19 Cullis, A. F., Muirhead, H., Perutz, M. F., Rossmann, M. G., and North, A. C. T., *Proc. R. Soc.*, **A265**, 15–38 (1961).
20 Dickerson, R. E., Weinzierl, J. E., and Palmer, R. A., *Acta Crystallogr.*, **B24**, 997–1003 (1968).
21 Blow, D. M., and Crick, F. H. C., *Acta Crystallogr.*, **12**, 794–802 (1959).
22 Debey, P., Balny, C., and Douzou, P., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 2633–2636 (1973).
23 Maurel, P., and Douzou, P., *J. biol. Chem.*, **250**, 2678–2680 (1975).
24 Fink, A. L., *Biochemistry*, **13**, 277–279 (1974).
25 Balny, C., and Hastings, J. W., *Biochemistry*, **14**, 4719–4723 (1975).
26 Chance, B., Saronio, C., and Leigh, J. S., Jr, *J. biol. Chem.*, **250**, 9226–9237 (1975).

letters to nature

Do freely falling bodies radiate?

IT seems clear that, although gravitational radiation has not yet been detected, its existence is predicted by general relativity. Whether freely falling bodies, that is, bodies moving under gravitation alone, radiate gravitational waves is, however, quite unclear. This is an important question because most astronomical motion is free fall, and if radiation is not produced by this, there will be much less of it present than some investigators believe. The situation has been clearly explained by Ehlers, Rosenblum, Goldberg and Havas (unpublished). Briefly, the doubts arise because those approximation methods which quite convincingly predict gravitational radiation from certain moving systems do not apply to bodies in free fall.

The question could be answered by suitable exact solutions of Einstein's equations. Until recently, however, the known exact solutions for freely falling matter have been spherically, plane or cylindrically symmetric or have been homogeneous: that is, they have had too much symmetry to be realistic models for radiating systems.

An exact solution has now been discovered[1] which is sufficiently unsymmetric to throw important light on the subject. It is a solution of Einstein's equations for dust, which is, of course, matter under no forces except gravity. The metric is

$$ds^2 = -e^{\lambda}dr^2 - e^{\omega}(dy^2 + dz^2) + dt^2 \qquad (1)$$

where

$$e^{\frac{1}{2}\omega} = \frac{\varphi(r,t)}{P(r,y,z)}; \quad e^{\frac{1}{2}\lambda} = \frac{P}{W(r)}\frac{\partial}{\partial r}e^{\frac{1}{2}\omega} \qquad (2)$$

$$P = a(r)(y^2 + z^2) + 2f(r)y + 2g(r)z + c(r) \qquad (3)$$

$$ac - f^2 - g^2 = \tfrac{1}{4} \qquad (4)$$

φ is a Friedmann function, that is, it satisfies the Friedmann equation

$$\dot{\varphi}^2 = [W(r)]^2 - 1 + \varphi^{-1}S(r) \qquad (5)$$

where the dot means $\partial/\partial t$; a, f, g, c, W, S are arbitrary subject to equation (4), and a further arbitrary function arises in the integration of equation (5). The density ρ is given by

$$8\pi\rho = \frac{PS' - 3SP'}{\varphi^2(P\varphi' - P'\varphi)} \qquad (6)$$

where the prime means $\partial/\partial r$. By a suitable choice of arbitrary functions the density may be made positive everywhere; and it can also be made to fall off with r as rapidly as one likes, so the solution may be thought of as referring to a cloud of dust concentrated towards the centre but extending very tenuously to infinity.

The motion of the dust, governed by equation (5), is one of expansion or collapse, as in the Friedmann cosmological models, but dependent on the radial coordinate r. The two-dimensional surfaces r = constant, t = constant are spheres and y, z are stereographic coordinates. The spheres of different radii are not concentric, however, and taken as a whole the three-dimensional space of r, y and z is not spherically symmetrical. In fact the solution has no Killing vectors whatever for general values of the arbitrary functions (W. B. B., A. H. Sulaiman and N. Tomimura, unpublished).

Does the space–time contain gravitational radiation? It is not easy to answer this question: there is no conclusive mathematical test for the presence of radiation in a space–time, and to examine the solution for a radiative flux at spatial infinity turns out to be difficult because of the comoving coordinate system in which the solution is formulated. The following argument seems conclusive, however, Consider the portion of the solution (1)–(6) contained within $0 \leqslant r < r_0$ (r_0 = constant). I have shown (unpublished) that it is possible to match this portion to a Schwarzschild metric which is valid for $r > r_0$. The Schwarzschild metric being static there is no radiation passing through it, so the moving dust in the region $0 \leqslant r < r_0$ is not radiating.

It would be quite wrong to argue from this that freely falling matter never radiates. The investigation does, however, show the existence of a class of non-radiative motions of dust in a space–time without Killing vector symmetries.

W. B. BONNOR

Department of Mathematics,
Queen Elizabeth College,
Campden Hill Road,
London W8, UK

[1] Szekeres, P., *Commun. math. Phys.*, **41**, 55–64 (1975); *Phys. Rev., D*, **12**, 2941–50 (1975).

Globular clusters as a source of X-ray emission from the neighbourhood of M87

AT least five galactic X-ray sources have been identified with globular clusters[1-3], and it is possible that a similar association holds for the bright unidentified sources near the galactic centre[4]. Intense X-ray bursts have been detected[5] from the direction of one of these (3U1820−30/NGC 6624), and a star cluster, probably globular, has been found[6] in the error box[7,8] of the so-called 'rapid burster', MXB1730−335. Moreover, the error box of a hard X-ray flare observed from the Cosmos 428 satellite[9] contains NGC5904 (M5). This globular cluster has similar properties[10] to the others that emit X rays, and if normally weak it is unlikely to have appeared in the Uhuru catalogue because of its position. The X-ray emission from globular clusters may be attributable to accretion on to compact objects, the accreting material being supplied from binary companions[11-14], or gas trapped in the potential well of the cluster[15,16].

There are relatively few intrinsically bright optically observed globular clusters in our Galaxy. Counts of objects in the vicinity of M87, however, reveal that it has an extensive halo of globular clusters[17,18], the number of which may exceed 10,000 within a radius of 23 arc min (ref. 19). The greater number of bright globular clusters in M87 may be explicable as a population effect[18]. The similarities

between the optical properties of these globular clusters and those in our own Galaxy suggest that the former may also contain X-ray sources.

The brighter globular clusters in M87 may also be substantially more X-ray luminous. The presence of a massive black hole or of captured binaries depends sensitively on the properties and past history of the core of the globular cluster. Gas to fuel a massive black hole, or single neutron stars, in a globular cluster orbiting our Galaxy will be stripped on passing through the galactic disk, and possibly even through the halo[20]. This stripping will not occur in an elliptical galaxy; consequently, there may be proportionally more gas available in identical globular clusters in M87 than in our Galaxy.

On either hypothesis, the average X-ray luminosity of individual globular clusters may be of the order of 10^{38} erg s^{-1}. This raises the possibility that the integrated globular cluster emission may account for a substantial fraction of the X-ray emission observed from the region of M87 ($L_X \sim 7 \times 10^{42}$ erg s^{-1}; ref. 21). (Katz[22] has suggested that the X-ray emission from clusters of galaxies may be attributable to unspecified compact sources.) In support of this proposition we note that the extended X-ray emission[21,23] from the Virgo cluster is centred on M87, which lies ~45 arc min from the cluster centroid. We expect that the general X-ray emission from the globular cluster will appear to be smoothly and symmetrically distributed about M87 at moderate spatial resolution. A few clusters of higher luminosity might be resolvable with currently proposed imaging X-ray telescopes.

A similar situation may apply to the elliptical galaxy NGC3311 in Abell 1060, which as a cluster has been suggested as the identification for the X-ray source 3U1044−30[1,24]. It seems possible that that galaxy is surrounded by a similar globular cluster population to that of M87 (ref. 25).

We thank Dr David Hanes for valuable discussions.

A. C. FABIAN
J. E. PRINGLE
M. J. REES

Institute of Astronomy,
Madingley Road,
Cambridge CB3 0HA, UK

1 Giacconi, R., *et al.*, *Astrophys. J. Suppl.*, **27**, 37–64 (1974).
2 Clark, G. W., Markert, T. H., and Li, F. K., *Astrophys. J. Lett.*, **199**, L93–L96 (1975).
3 Canizares, C. R., and Neighbours, J. E., *Astrophys. J. Lett.*, **199**, L97–L99 (1975).
4 Canizares, C. R., *Astrophys. J.*, **201**, 589–592 (1975).
5 Grindlay, J., Schreier, E., Schnopper, H., Gursky, H., Parsignault, D., and Heise, J., *Astrophys. J. Lett.*, **205**, L127–L130 (1976).
6 Liller, W., *IAU Circ.*, No. 2928.
7 Lewin, W. H. G., *et al.*, *Astrophys. J. Lett.*, **207**, L95–L100 (1976).
8 Heise, J., *et al.*, *Nature*, **261**, 562–564 (1976).
9 Babushkina, O. P., *et al.*, *Lett. Sov. Astr.-A.J.*, **2**, 20; **6**, 6 (1975).
10 Peterson, C. J., and King, I. R., *Astr. J.*, **80**, 427–436 (1975).
11 Clark, G. W., *Astrophys. J.*, **199**, L143–L145 (1975).
12 Fabian, A. C., Pringle, J. E., and Rees, M. J., *Mon. Not. R. astr. Soc.*, **172**, 15P–18P (1975).
13 Sutantyo, W., *Astr. Astrophys.*, **44**, 227–230 (1975).
14 Hills, J. G., *Mon. Not. R. astr. Soc.*, **175**, 1P–3P (1976).
15 Bahcall, J. N., and Ostriker, J. P., *Nature*, **256**, 23–24 (1975).
16 Silk, J., and Arons, J., *Astrophys. J. Lett.*, **200**, L131–L135 (1975).
17 Sandage, A., *Hubble Atlas of Galaxies* (1961).
18 Hanes, D., thesis, Univ. Toronto (1975).
19 Harris, W. E., and Smith, M. G., *BAAS*, **7**, 534 (1975); *Astrophys. J.*, **207**, 1036–1045 (1976).
20 Frank, J., and Gisler, G., *Mon. Not. R. astr. Soc.* (in the press).
21 Kellogg, E., in *X-ray Astronomy*, (edit. by Giacconi, R., and Gursky, H.) (Reidel Dordrecht, 1975).
22 Katz, J. I., *BAAS*, **7**, 539 (1975); *Astrophys. J.*, **207**, 25–28 (1976).
23 Malina, R., Lampton, M., and Bowyer, S., *Astrophys. J.* (in the press).
24 Ives, J. C., and Sanford, P. W., *Mon. Not. R. astr. Soc.*, **176**, 13P–17P (1976).
25 Smith, M. G., and Weedman, D. W., *Astrophys. J.*, **205**, 709–715 (1976).

Pre-terrestrial shear faulting and heat treatment of the Jamestown iron meteorite

WE have made a metallographic examination of a 250-g slice of the Jamestown iron meteorite. The meteorite was first reported by Huntington[1] as a mass of 4 kg, with a dished,

flake-like shape of maximum dimensions $26 \times 13 \times 3.7$ cm, and with two entirely different external surfaces—a smooth convex side, and a slaggy, vesicular, concave side, pitted by circular cavities up to 2 cm in diameter. Buchwald[2] has reported on the general metallography of the Jamestown Iron but does not record which of the many sections of this meteorite his description is based on.

Our British Museum specimen (BM 67215) mates with that shown in Fig. 3 of ref. 1. Most of our general microscopic observations are in agreement with those of Buchwald[2]. We found that phosphide is absent, daubreelite is present as rare particles, and troilite is present in varying degrees of shock alteration. The size and shape of the ε-kamacite, taenite and plessite components are consistent with membership of chemical group IVA in the classification scheme of iron meteorites.

Furthermore, the convex surface of the British Museum specimen retains relics of a heat alteration zone, produced by ablative burning during atmospheric entry. The concave surface shows somewhat less penetration of heat alteration but instead shows a considerable accumulation of the products of ablation melting, which Buchwald[2] describes as fine grained metallic eutectics. The Jamestown meteorite must have traversed the Earth's atmosphere as an oriented projectile with the convex surface leading.

The surface heat alteration overlies all other aspects of metallic macro- and microstructure which must therefore be of pre-terrestrial origin and not produced by impact with the surface of the Earth. Buchwald[2] has noted that the ε-kamacite must have been reheated. Though it is not easy to specify the exact conditions of heat treatment, some limits may be proposed, since on low power examination the kamacite seems to consist of the acicular ε structure noted by Jain and Lipschutz[3], but our examination at higher magnification shows that the ε within the bulk of the kamacite is diffused but not recrystallised. This absence of recrystallisation in the bulk of the kamacite suggests a reheating temperature below 450 °C, since our laboratory heat treatment of the shock-hardened kamacite of Trenton does show detectable recrystallisation at this temperature.

Buchwald[2] noted local distortion, shearing and folding of plessite and taenite, but did not indicate the scale of these effects and did not examine the detailed metallography of the shearing, which in some instances is encountered on a macroscopic, not a local, scale. Indeed, the outstanding feature of the etched macrostructure of the British Museum specimen is the interruption to the regularity of the Widmanstätten pattern by the traces of at least five shear-displacement surfaces, of which the longest extends for 6 cm and leaves the meteorite at two points on the leading convex edge. Others extend from the trailing concave surface inwards for about 2 cm. One of the latter has developed into the shear-displacement type of crack that Axon and Steele-Perkins[4] have reported in fragments of the Henbury Iron. It is, however, important to recognise that the shear-displacements in the Henbury specimen were produced by the crater-forming impact of the meteorite with the Earth's surface; but the shear-displacement crack in the Jamestown Iron is invaded by ablation melt product and is therefore of pre-terrestrial origin.

Attention has been directed to the absence of recrystallation in the bulk of the ε-kamacite, and a reheating temperature < 450 °C deduced. By contrast, the more heavily deformed material within and immediately adjacent to the surfaces of shear displacement in the Jamestown Iron has recrystallised. The zones of finely recrystallised kamacite extend on both sides of the shear surface to a width of about 0.15 mm, and it is this band of recrystallisation that accentuates the appearance of shear surfaces in the macrostructure. This local recrystallisation of the heavily sheared material is consistent with the proposed reheating temperature, and explains why the macroscopically visible thickness of the shear traces (~ 0.15 mm) in the Jamestown Iron is greater than in other meteorites previously examined by us.

The metallographic features of the Jamestown meteorite would be consistent with the following history. After a period of slow cooling (approximately 300 °C Myr^{-1}, according to the bandwidths reported by Buchwald[2]) in which the Widmanstätten structure developed, the material formed part of a massive projectile that impacted, and formed a crater on, some unknown planetismal in the Solar System. The present Jamestown meteorite is part of a fragment that was produced during that crater-forming event. It shows shear displacement surfaces and the development of fracture by separation along the surfaces of shear faults in a manner similar to that noted by Axon and Steel-Perkins[4] for the terrestrial impact of the Henbury Iron. After the crater-forming impact the Jamestown material was heated to a temperature below 450 °C, probably by a blanket of hot ejecta on the surface of the unknown planetismal. At some later stage it was released from its intermediate host, circulated in space and eventually entered the Earth's atmosphere, where it acquired its heat alteration zone and an accumulation of ablation deposit, some of which penetrated a pre-existing shear fissure on the rear surface of the oriented projectile.

Shock effects are common in group IVA irons, and Jain and Lipschutz[3] have noted the presence of reheating effects after shock in the Boogaldi, Western Arkansas, Huizopa, Charlotte, Yanhuitlan, Maria Elena and Social Circle meteorites. The massive shear traces in the Jamestown Iron are, however, so far unique within the group IVA irons, although the Seneca Township meteorite shows a similarly diffused ε structure in its bulk.

We thank John Wasson, who provided early copies of ref. 2.

H. J. Axon

Metallurgy Department,
University of Manchester,
Manchester M1 7HS, UK

A. W. R. Bevan

Mineralogy Department,
British Museum (Natural History),
Cromwell Road,
London SW7 5BD, UK

1 Huntington, O. W., *Proc. Am. Acad. Arts Sci.*, 25, 229 (1890).
2 Buchwald, V. F., *Handbook of Iron Meteorites*, 690–691 (Univ. California Press, 1976).
3 Jain, A. V., and Lipschutz, M. E., *Geochim. cosmochim. Acta*, 34, 883 (1970).
4 Axon, H. J., and Steele-Perkins, E. M., *Nature*, 256, 635 (1975).

Lead and radium in the lower stratosphere

More than a decade ago Junge[1] found that the concentration of particles in the stratosphere is greater than in the upper troposphere, and that their main chemical constituents were ammonium and sodium sulphates, thought to arise from the ascent of their volatile precursors H_2S and SO_2 (ref. 2). In addition, minute concentrations of aluminium, calcium, chlorine, chromium, cobalt, copper, gold, iron, manganese, silicon and titanium have been detected[3,4], which, having no volatile precursors, enter the stratosphere as particles. As the extraterrestrial fraction in stratospheric aerosols is extremely small[5], their main source seems to be the surface of the Earth, but data available on the nature of such sources, and on the chemical composition and concentration of stratospheric aerosols are rather scarce—an astonishing fact, in view of their possible climatic impact. With this in mind, we have measured the concentrations of stable lead, ^{226}Ra and its radioactive daughter ^{210}Pb at various heights. These substances enter the atmosphere from several surface sources, both natural and artificial[6]. The concentration of ^{226}Ra in particulate materials such as soil dust or fly ash is well known, and may be used for the assessment of the stratospheric content of this kind of dust.

During the past three years we collected 112 samples of
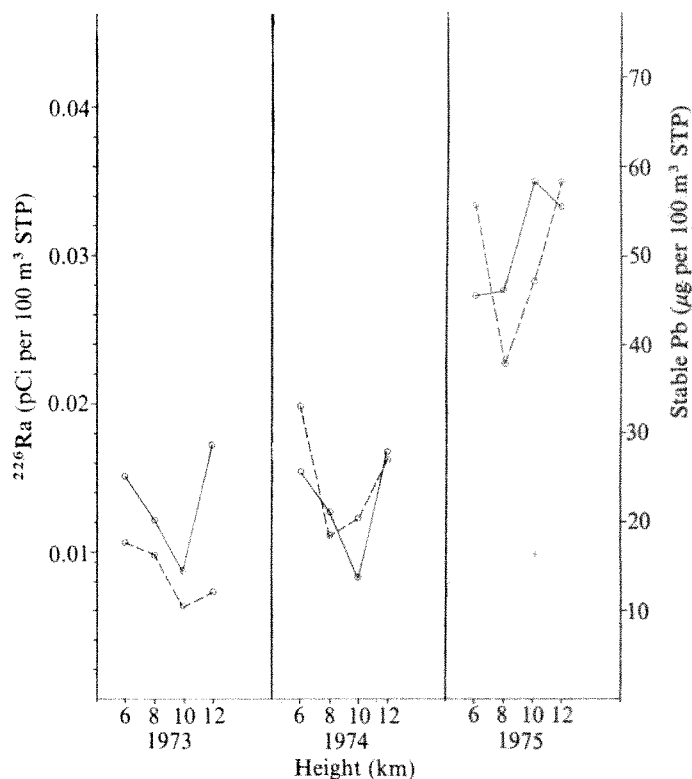
Fig. 1 Vertical distribution of ^{226}Ra (– – – –) and stable Pb (—) in the atmosphere.

aerosols over Poland, from heights of 6, 8, 10 and 12 km. At each altitude 9–15 samples were collected in 1973, 6–9 in 1974 and 6–11 samples in 1975. The levels of the tropopause for each sampling profile were determined from data on the vertical gradient of temperature, supplied by the Institute of Hydrometeorology, Warsaw.

We used 1,600-cm^2 PVC fibre filters, type FPP-15-1.7 (Soviet made), placed in collectors suspended under the wings of LIM-type jet planes. After wet-ashing and dissolving the filters, the quantities of ^{226}Ra, stable Pb, ^{210}Pb, ^{90}Sr and ^{137}Cs collected were determined, using methods described elsewhere[7].

In Fig. 1 we present the annual mean air concentrations of stable Pb and ^{226}Ra, found between July and November 1973, May and October 1974 and March and December 1975. The striking feature is the high concentration (comparable with ground level) of ^{226}Ra and stable Pb at 12 km. In 1973 the mean ^{226}Ra concentration at 12 km was 0.7×10^{-4} pCi m^{-3} STP, 1.6×10^{-4} pCi m^{-3} STP in 1974 and 3.5×10^{-4} pCi m^{-3} STP in 1975. The mean stratospheric stable-Pb concentrations were 29×10^{-2} µg m^{-3} STP in 1973 and 1974 and 55.7×10^{-2} µg m^{-3} STP in 1975. The concentrations of ^{226}Ra which we found in Warsaw in the air at ground level range from 1.5×10^{-4} to 6.0×10^{-4} pCi m^{-3} and with the mean stable-Pb concentration of 10.0×10^{-2} µg m^{-3} in non-urban air[8].

Comparing Figs 1 and 2, it is also interesting to note that the vertical gradients of ^{226}Ra and stable-Pb concentrations in the troposphere are opposite to the gradients of the fission products. The concentrations of ^{226}Ra and Pb decrease with altitude between 6 and 10 km, reaching minima below the tropopause, and increasing sharply above it. This shows the ground level origin of these substances, and indicates the existence of trapping in the stratosphere.

Stable Pb, ^{226}Ra and ^{210}Pb have been discharged into the atmosphere by natural sources since time immemorial. Volcanic activity, forest fires and the action of the wind on soil and sea water are estimated to contribute ~ 5.4×10^{-4} µg of Pb m^{-3} of atmospheric air[9] and 3×10^{-9} pCi m^{-3} of ^{226}Ra (ref. 6). But the present concentrations of stable Pb and ^{226}Ra in the stratosphere are 2–3 and 5 orders of magnitude higher than these estimated natural levels. It follows that the bulk of
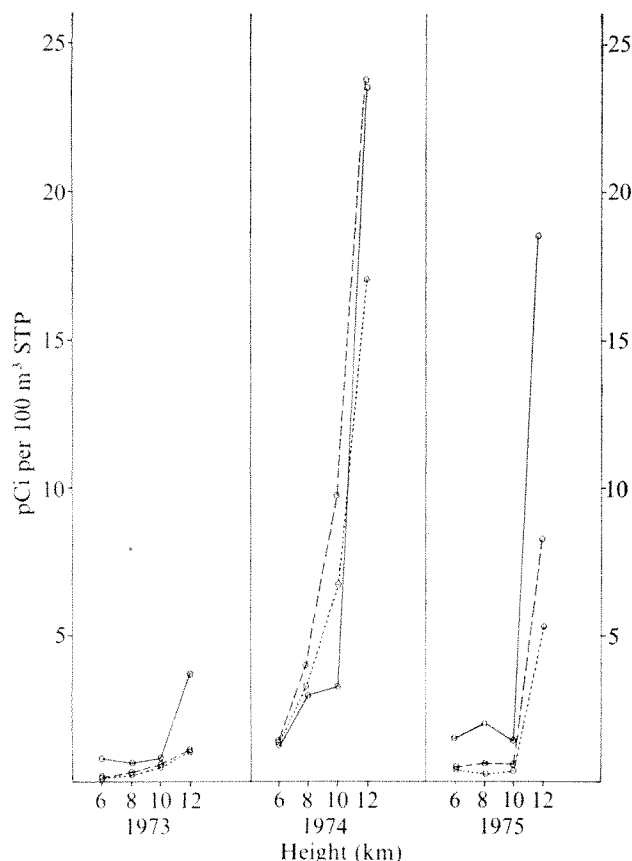
Fig. 2 Vertical distribution of ^{210}Pb (—), ^{137}Cs (- - - -) and ^{90}Sr (· · · · ·) in the atmosphere.

these nuclides in the stratosphere is probably of artificial origin, unless the estimation of their natural input is wrong. Coal burning, phosphate fertilisers and cement production are the main man-made sources of ^{226}Ra in the atmosphere; their global contributions have been calculated as 150, 400 and 1.8 Ci yr^{-1} respectively[6].

If the mean concentration of ^{226}Ra in the $\sim 10^{19}$ m^3 of atmosphere up to 20 km, were similar to that found in the lower stratosphere, that is $\sim 1.0 \times 10^{-4}$ pCi m^{-3}, then the total content of ^{226}Ra in this volume would be $\sim 1,000$ Ci, a value not much different than the annual contribution from artificial sources.

The main artificial sources of stable Pb are the combustion of lead alkyls in petrol, which contributes 120,000 t yr^{-1} (Northern Hemisphere)[10] and the burning of coal, contributing 3,000 t yr^{-1} (whole world)[8]. For the same volume of atmosphere, with a mean Pb concentration of 1.0×10^{-2} μg m^{-3}, the total lead content would be 100,000 t. These tentative calculations support the suggestion that most of ^{226}Ra and Pb in the stratosphere is of artificial origin.

During the years 1973 and 1974 a series of Chinese and French nuclear explosions has been conducted in the atmosphere, which is responsible for the increase of ^{137}Cs and ^{90}Sr in our samples in 1974 and 1975. It is interesting to see in Fig. 2 that ^{210}Pb concentrations in 1974 increased in parallel with those of fission products. In the first half of 1975, ^{210}Pb concentrations increased above the 1974 level, whereas a substantial decrease of ^{137}Cs and ^{90}Sr was observed. At the same time we found a sharp increase in the stable Pb content in both troposphere and stratosphere, seeming to suggest a common source for both types of lead. This is, however, improbable, as the specific activity of lead collected by us in the lower stratosphere is extremely high, reaching, in 1975, $\sim 330,000$ pCi ^{210}Pb per g stable lead, and, in 1974, $\sim 2,400,000$ pCi g^{-1}, 3–6 orders of magnitude above the specific activities of contemporary common lead, which range from 1.5 to 501 pCi g^{-1}

(ref. 12). This indicates that common Pb in the atmosphere is of no importance as a source of ^{210}Pb.

From the known concentrations of ^{226}Ra and stable Pb in coal[8,13], one can infer that the specific activity of lead in coal is $\sim 30,000$ pCi per g lead. In soil this specific activity is 15,000 pCi per g lead[11]. It follows that although coal and soil might contribute to stratospheric load of ^{210}Pb, the bulk of it comes from other sources. Coal burning contributes yearly ~ 200 Ci, and phosphate fertilisers 150 Ci (ref. 6). Natural production of ^{210}Pb in the atmosphere from ^{222}Rn, estimated as ~ 0.62 MCi yr^{-1} (ref. 12), clearly overwhelms these sources. Nuclear explosions are suggested to have produced fractions of a MCi of ^{210}Pb during the 1958–59 and 1961–62 test series, by the reaction ^{208}Pb(2n,γ)^{210}Pb, from the lead content of the construction of bombs[14,15]. This contribution, being transient and of short duration, would explain the excess peak concentrations of ^{210}Pb in the stratosphere accompanied by high concentrations of fission products as observed by us and others (P. W. Krey, unpublished). It also explains similar correlations found in fossil ice collected at glaciers in the Tatra Mts, Poland, the Jotunheimen Mts, Norway and the Himalayas, Nepal[11].

The minute industrial sources of ^{210}Pb, cited above, may locally influence the content of this nuclide in air and precipitation, as suggested by several authors[6,14,16]. But the contributions from nuclear explosions, could affect the ^{210}Pb concentration on a global scale. This restricts the use of ^{210}Pb as a tracer in determinations of the age of glacier ice and in atmospheric transport studies.

If the ^{226}Ra content in stratospheric particulates were the same as in the fly ash or soil, ~ 1–5 pCi g^{-1} (ref. 11), then the 1975 concentration of mineral matter bound to ^{226}Ra would reach 50–250 μg m^{-3} STP in stratospheric air, at a height of 12 km. This is a rather high concentration. Were these particulates of artificial origin, it would imply that their accumulation in the stratosphere is a fairly new process, which if continued indefinitely might have an important climatic impact in the future.

This work has been carried out under the auspices of the US Environmental Protection Agency.

Z. JAWOROWSKI
L. KOWNACKA

Central Laboratory for Radiological Protection,
Warsaw, Poland

1 Junge, C. E., *Air Chemistry* (Academic, New York, 1963).
2 Junge, C. E., and Manson, J. E., *J. geophys. Res.*, **66**, 2163–2182 (1961).
3 Shedlovsky, J. P., and Paisley, S., *Tellus*, **18**, 498–503 (1966).
4 Gillette, D. A., and Blifford, J. H., Jr, *J. atmos. Sci.*, **28**, 1199–1210 (1971).
5 Ferry, G. V., and Lem, H. Y., *Climatic Impact Assessment Program* (edit. by Broderick, A. J., and Hard, T. M.), 310–317 (US Department of Transportation, Washington, DC, Report No. DOT-TSC-OST-74-15, 1974).
6 Jaworowski, Z., Bilkiewicz, J., Kownacka, L., and Wlodek, S., *The Natural Radiation Environment II* (edit. by Adams, J. A. S., Lowder, W. M., and Gesell, T. F.), 809–818 (USERDA, Report CONF-720805-P2, 1972).
7 Jaworowski, Z., Bilkiewicz, J., and Dobosz, E., *Environ. Pollut.*, **9**, 305–315 (1975).
8 Jaworowski, Z., *Stable and Radioactive Lead in Environment and in Human Body*, *Review Report*, **29**, NEIC-RR-29, 1–179 (Nuclear Energy Information Center, Warsaw, 1967).
9 Patterson, C. C., *Arch. Environ. Health*, **11**, 344–363 (1965).
10 Murozumi, M., Chow, T. J., and Patterson, C. C., *Geochim. cosmochim. Acta*, **33**, 1247–1291 (1969).
11 Jaworowski, Z., *et al.*, *The Biological Implications of Metals in the Environment —Proc. 15th Hanford Life Sciences Symp. 1975* (USERDA Symposium Series, in the press).
12 Jaworowski, Z., *At. Energy Rev.*, 7, 3–45 (1969).
13 Jaworowski, Z., and Bilkiewicz, J., *Health Phys.*, **20**, 449–450 (1970).
14 Peirson, D. H., Cambray, R. S., and Spicer, G. S., *Tellus*, **18**, 427–433 (1966).
15 Jaworowski, Z., *Nature*, **212**, 886–889 (1966).
16 Bonnyman, J., Duggleby, J. C., and Molina-Ramos, J., *The Natural Radiation Environment, II* (edit. by Adams, J. A. S., Lowder, W. M., and Gesell, T. F.), 819–831 (USERDA, Report CONF-720805-P2, 1972).

Progressive faunal migration across the Iapetus Ocean

DURING the Lower Palaeozoic, there was a gradual increase in the similarity of the faunas between North America to the west of the Appalachians, western Newfoundland, northwestern Ireland and Scotland on the one hand, and coastal
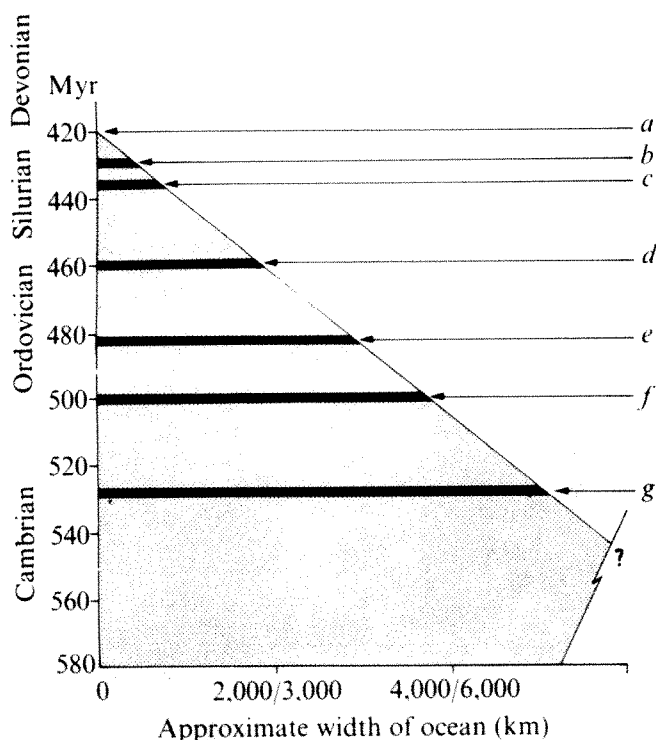
Fig. 1 The times at which faunas became common to both sides of the Lower Palaeozoic Iapetus Ocean. *a*, Closure of ocean (Norway); *b*, freshwater fish; *c*, benthic ostracods; *d*, trilobite and brachiopod species; *e*, trilobite and brachiopod genera (? island hopping); *f*, *Didymograptus bifidus*; *g*, *Dictyonema*.

New England, southern New Brunswick, Nova Scotia, eastern Newfoundland, England and northern Europe on the other hand. The best explanation for this is that the Iapetus (or Proto-Atlantic) Ocean was wide enough to separate two faunal provinces in Cambrian times, and that there was progressive migration of the more mobile components of the faunas as the old ocean closed (Fig. 1). The pelagic animals crossed first, followed later by animals (trilobites and brachiopods) with pelagic larval stages; but animals without a pelagic larval stage (such as benthic ostracods) were not able to cross until the ocean had closed at one point, though not necessarily everywhere along its length. Finally, faunas limited to fresh water or brackish water (like many Devonian fish) did not cross until there were non-marine connections between the continents on either side of the closing ocean.

Pelagic graptolites were among the earliest animals to appear on both sides of the ocean. The many-branched *Dictyonema flabelliforme* migrated across in the Tremadoc; but it was not until the Llanvirn (towards the end of the early Ordovician) that most other graptolites became similar on both sides. The two-branched *Didymograptus bifidus*, for example, appears in the Arenig of Texas, but did not reach England and Wales until the Llanvirn[1].

The next groups of animals to cross were those with pelagic larval stages, notably brachiopods and trilobites. Most genera in these groups were confined to one or other side of the Iapetus Ocean during the Cambrian and early Ordovician; the few cases where genera characteristic of different faunal provinces occurred together (including New World Island, Newfoundland) have been interpreted[2] as sites of oceanic islands colonised by sporadic spatfalls, which became centres of evolution and migration. The late Ordovician (Caradoc) brachiopod and trilobite faunas show less marked distinctions[3,4]; though most species are distinct, many genera are common to both sides of the ocean. It is not until the latest Ordovician (latest Ashgill; post-Richmond) that most species of brachiopods and trilobites became identical in Europe and America.

The specific distinctions in the Caradoc and early Ashgill indicate that, at its narrowest point, the Iapetus Ocean must have still been wide enough to prevent free crossing by most trilobite and brachiopod larvae. There are two quite distinct ways of measuring the width of the ocean at this time: by analogy with modern pelagic larva, and by analogy with modern rates of subduction.

Most modern brachiopod larvae have a pelagic stage of less than one week, though it may be as long as three weeks[5], and 80% of decapod larvae have ranges of between 2 and 7 weeks[6]. But benthic marine invertebrates commonly have a much longer larval life in the tropics, (for example, 7 to 43 weeks for tropical gastropods[7]). Britain and eastern Canada were at about 20°–35°S during the late Ordovician[8], so we conclude that the late Ordovician brachiopod and trilobite larvae (which seem to have crossed the Iapetus Ocean more or less simultaneously) may have had a larval duration of about 7 to 14 weeks. Data from Thorson[6] and Scheltema[7] suggest that larvae with this lifespan could possibly travel across an ocean between 2,000 and 4,000 km wide.

We have estimated that subduction occurred to the north of the Iapetus Ocean (in Scotland and north-western Ireland) between 530 (the age of the Oughterard granite of County Galway) and 390 Myr ago (the Glencoe Caldera of Scotland), a duration of 140 Myr; southward subduction occurred between 490 Myr ago (the Borrowdale Volcanics of northern England) and 440 Myr ago (the Mendip Hills and the Skomer Volcanics). If the average subduction rate during these periods was 4 cm yr^{-1}, then some 7,600 km of ocean crust may have been subducted between the late Cambrian and mid Silurian. Of this, some 4,800 km would have been subducted before the late Ashgill; thus, some 2,800 km of ocean crust might have existed in the late Ashgill. This argument assumes that an active spreading centre no longer existed in the late Ashgillian Iapetus Ocean; if it did exist the ocean could have been narrower. The estimates from pelagic larval migration and from subduction rates thus both suggest that the ocean had a minimum width of around 2,000–3,000 km during the late Ashgill.

The Iapetus Ocean probably did not have parallel margins; there is some evidence to suggest that it may have been wider in eastern Canada and the northern Appalachians than in Europe. The first closure occurred in the north-east[9], when Greenland and the Baltic Shield collided during the late Silurian. This date is supported by evidence from freshwater fish : the first recorded species to cross the Iapetus Ocean did so between Scotland and Norway during Wenlock or Ludlow times[10].

Subduction of ocean floor continued through the Silurian and early Devonian in the region between Britain and the northern Appalachians, and continental collision in this region (as opposed to Greenland and Norway) did not take place until after a large proportion of the Lower Devonian rocks had formed. The time of collision in this area is clearly related to the Acadian Orogeny, which can be dated as middle Emsian, that is, towards the end of early Devonian time[11].

W. S. McKerrow

Department of Geology and Mineralogy,
Parks Road,
Oxford OX1 3PR, UK

L. R. M. Cocks

Department of Palaeontology,
British Museum (Natural History),
Cromwell Road,
London SW7 5BD, UK

[1] Bergstrom, S. M., and Cooper, R. A., *Lethaia*, **6**, 313–339 (1973).
[2] Neuman, R. B., *Proc. 24th int. Geol. Congr., Montreal*, **7**, 297–302 (1972).
[3] Williams, A., in *Spec. Pap. Palaeont.*, **12**, 241–269 (1973).

4 Whittington, H. B., in *Atlas of Palaeobiology* (edit. by Hallam, A.), 13–18 (Elsevier, Amsterdam, 1973).
5 Rudwick, M. J. S., *Living and Fossil Brachiopods* (Hutchinson, London, 1970).
6 Scheltema, R. S., *Biol. Bull.*, **140**, 284–322 (1971).
7 Thorson, G., in *Oceanography* (edit. by Sears, M.), *Am. Ass. Advan. Sci.*, **67**, 455–474 (1961).
8 Briden, J. C., Drewry, G. E., and Gilbert Smith, A., *J. Geol.*, **82**, 555–574 (1974).
9 Dewey, J. F., and Kidd, W. S. F., *Geology*, **2**, 543–546 (1974).
10 Turner, P., and Turner, S., *Norsk. geol. Tidds.*, **54**, 183–192 (1974).
11 Boucot, A. J., *et al.*, *Spec. Pap. geol. Soc. Am.*, **139**, (1974).

Chronological evolution of the Kerguelen Islands syenite–granite ring complex

ALKALINE syenites in the Rallier-du-Baty peninsula of the Kerguelen Islands[1,2] intrude and metamorphose the basaltic lava flows which constitute the major part of the islands[3]. Recent work[4] has revealed the existence, in the peninsula, of five ring complexes consisting mainly of syenites containing from 2 to 15% normative quartz (Figs 1 and 2). The southern complex is the largest, and consists of nordmarkites (ring dykes *a, b, c*) and of quartz-rich syenites and alkaline granite (ring dykes *d, e, f*). Some minor intrusive bodies of gabbros (G) and nordmarkites (Σ) are located on the margins of the southern centre (Fig. 2). Chilled margins and enclaves in each individual ring dyke clearly indicate that the sequence of intrusions is centripetal. Furthermore, the contact relationships between the centres show that the magmatic activity has migrated northwards. Previous K–Ar data[5] have indicated that a syenite from the southern ring complex was emplaced 8.7 ± 0.9 Myr ago, whereas basaltic flows were erupted during the late Oligocene. An additional K–Ar age of 11.5 ± 0.2 Myr was obtained for a
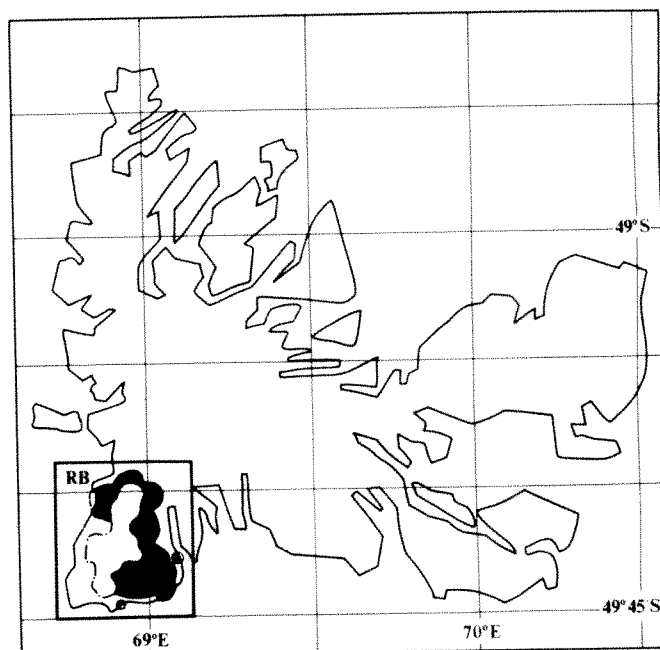
Fig. 1 Location map of the Rallier-du-Baty peninsula (RB) in the Kerguelen Islands.

metabasalt collected near the contact with the syenites and was interpreted[3] as the minimum value for the time of eruption of the basalts.

Using both K–Ar and Rb–Sr geochronological techniques, we have dated a series of samples, carefully selected from
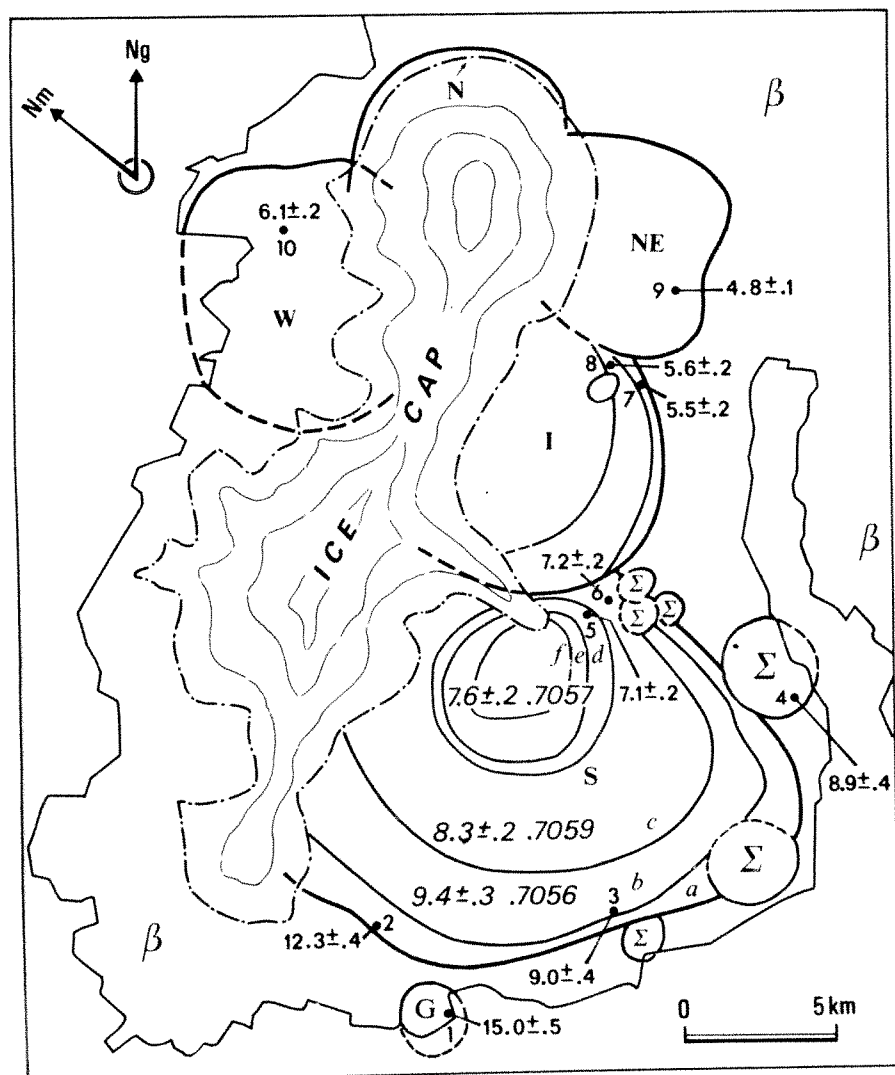
Fig. 2 Geological sketch map of the Rallier-du-Baty peninsula, K–Ar (J.M.C.) and Rb–Sr (L.D., Ph.V.) results. S, Southern centre; I, middle centre; NE, north-eastern centre; N, northern centre; W, western centre. K–Ar results: 1, olivine gabbro; 2, quartz–syenite; 3, aplite dyke; 4, fayalite syenite; 5, granite (chilled margin); 6, microgabbro; 7, quartz–syenite (chilled margin); 8, microgabbro (dyke); 9, syenite; 10, biotite syenite. Rb–Sr results (see Fig. 3). Nm, Magnetic North; Ng, geographical North.

Fig. 3 *b*, Whole-rock isochron for syenitic ring dyke *b* (○), 9.4±0.3 Myr; *c*, whole-rock isochron for syenitic ring dyke *c* (●) 8.3±0.2 Myr; *d, e, f,* whole-rock isochron for granitic and syenitic inner ring dykes (▲), 7.6±0.2 Myr. I and II, Duplicate analyses on certain samples.
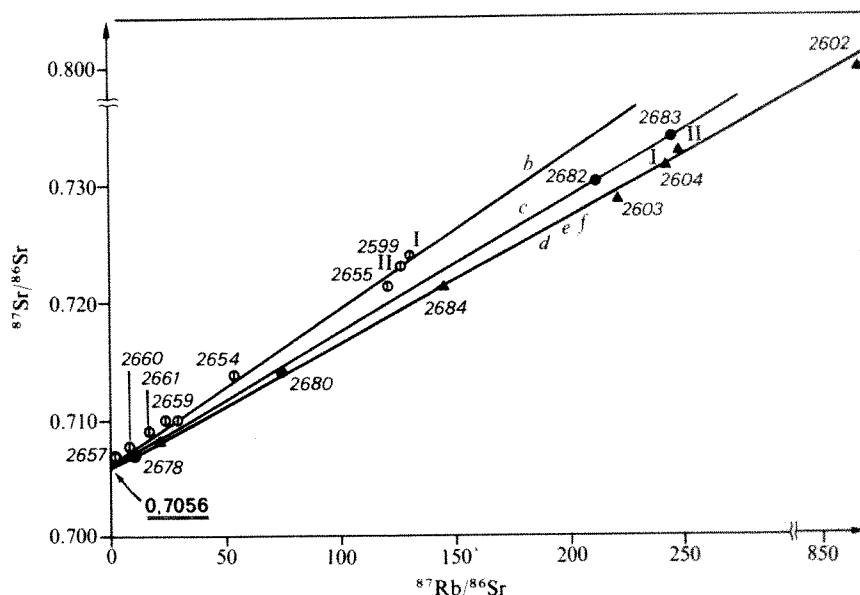
individual ring dykes in the southern centre (see Fig. 2 for sample localities) and have obtained the following results: (1) Both K–Ar and Rb–Sr methods yield consistent ages ranging from ∼ 15 Myr for the minor olivine–gabbro intrusion, and 12 Myr for the outer syenitic ring dyke, to ∼ 7 Myr for the inner ones and the core, with an interval of ∼ 1 Myr between the emplacement of adjacent ring dykes. (2) Most of the dated samples (Fig. 3) have very high Rb/Sr ratios (up to 300). Three whole-rock Rb–Sr isochrons yielded significantly different ages (9.4±0.3 Myr; 8.3±0.2 Myr; 7.6±0.2 Myr) but they all have very similar initial $^{87}Sr/^{86}Sr$ ratios. The 7.6±0.2 Myr isochron, however, combines various samples from the three inner distinct ring dykes (*d, e, f*) made up of alkaline granites and quartz syenites.

Some additional samples from the western, north-eastern, and intermediate centres have also been measured using the K–Ar method. The ages of the group of complexes range from > 12 Myr in the south, to < 5 Myr in the north-east. These data are in agreement with the deductions made from the field observations and suggest that the magmatic activity migrated at a very slow rate during the Upper Miocene.

The very high, yet variable, Rb/Sr ratios, as well as the radiogenic nature of the samples from the southern centre, allow a fairly precise whole-rock Rb–Sr isochron to be obtained for these very young rocks for the first time. The centripetal age distribution in the southern centre may reflect the consequence of a progressive cooling from the outer ring dyke to the core; however, the presence of chilled margins in the outer part of the ring dykes *a, b, c* and *d* (Fig. 2) practically rules out such a possibility. The whole-rock Rb–Sr ages are therefore interpreted as the individual times of crystallisation. The duration of about 5 Myr for the igneous activities in this ring complex is especially informative in terms of the formation of one complete ring complex. Although other similar ring complexes have already been studied in New Hampshire, Nigeria, Niger and Rhodesia[6–9], the actual time span for the development of igneous activity in a single ring complex could not be obtained because of the overlap of the analytical errors for the individual ages in successive intrusions.

The initial $^{87}Sr/^{86}Sr$ ratios for the three isochrons are: 0.7059±0.0003 for ring dyke *b*, 0.7056±0.0004 for ring dyke *c*, and 0.7057±0.0004 for all three inner ring dykes *d, e* and *f* and these values are in agreement with previous data[10,11]. Using Hedge's measurements, the calculated initial $^{87}Sr/^{86}Sr$ ratio is 0.7055 for three syenitic samples from the southern complex and one sample from a gabbroic intrusion of the Courbet peninsula. We suggest therefore, that the source regions for different ring dykes have the same Sr isotopic composition. Furthermore, the interval between the production of melt and the final crystallisation may have been very short (< 1 Myr in

this case), because, in our opinion, the very high Rb/Sr ratio could produce a very rapid increment in the $^{87}Sr/^{86}Sr$ ratio in the melt if fractional crystallisation was prolonged. The distinctly high initial $^{87}Sr/^{86}Sr$ ratios reported[7] in some geochemically similar rocks (Rb/Sr up to 1,200) from Nigeria which have previously been interpreted in terms of crustal contamination, may have an alternative interpretation if the liquids took a longer time to crystallise finally.

The question of whether continental crust or oceanic crust[10,11] exists beneath the Kerguelen Islands cannot be resolved from the Sr isotopic data alone. But neither do we have any geochemical or geological (xenolithic) evidence for the presence of continental material under the basalts.

J. LAMEYRE
A. MAROT
S. ZIMINE

Université de Paris VI,
Laboratoire de Pétrologie,
ERA 381 (CNRS),
4, Pl. Jussieu,
75230 Paris, France

J. M. CANTAGREL

Chronologie des Terrains
  Cristallins et Volcanologie, (CNRS),
5, Rue Kessler,
63, Clermont Ferrand, France

L. DOSSO
Ph. VIDAL

Centre Armoricain d'Etudes
  structurales des Socles (CNRS),
Ave du Gl. Leclerc,
35031 Rennes, France

1. Lacroix, A., *C.r. hebd. Séanc. Acad. Sci., Paris*, **179**, 113–119 (1924).
2. Aubert de la Rue, E., *C.r. hebd. Séanc. Acad. Sci., Paris*, **193**, 863–866 (1931).
3. Nougier, J., *Bull. Soc. Geol. Fr.*, **8**, 567–576 (1966).
4. Marot, A., Zimine, S., and Lameyre, J., *3rd Reunion Annuelle des Sciences de la Terre, Montpellier*, **248**, 395 (1975).
5. Nougier, J., *Com. natn. Francais Recherches Antarctiques*, no. 27, 440 (1969).
6. Foland, K. A., Quinn, A. W., and Giletti, B. J., *Am. J. Sci.*, **270**, 321–330 (1971).
7. van Breemen, O., and Bowden, P., *Nature phys. Sci.*, **242**, 9–11 (1973).
8. van Breemen, O., Hutchinson, J., and Bowden, P., *Contr. Miner. Petrol.*, **50**, 157–172 (1975).
9. Foland, K. A., and Henderson, C. M. B., *Earth planet. Sci. Lett.*, **29**, 291–301 (1976).
10. Nougier, J., and Lameyre, J., *Trans. Am. Geophys. Un.*, **53**, 547 (1972).
11. Watkins, N. D., Gunn, B. M., Nougier, J., and Baksi, A. K., *Bull. geol. Soc. Am.*, **85**, 201–212 (1974).

Native copper in DSDP sediment cores from the Angola Basin

LEG 40 of the Deep Sea Drilling Project (DSDP) drilled two holes in the Angola Basin: site 364 and site 365 (Fig. 1). Native copper was found in two samples (364-5-2, 19–21 cm and 364-5-2, 22–24 cm) from one of the cores taken at site 364. The botryoidal form of most of the copper grains initially suggested formation *in situ* and the subsequent discovery of copper casts of planktonic foraminifera removed any lingering thoughts of contamination. This is not the first report of native copper in DSDP cores. Most other occurrences have, however, been found in basement rocks (for example, site 282) or in the sediments immediately overlying basement rocks (site 105), whence the copper was presumably derived. The copper reported here occurs at least 3 km above volcanic basement rocks, and no copper has yet been found in samples above or below it. Only at site 149 (Venezuelan Basin) has copper previously been found in non-basement-associated sediments. There it occurs as "a number of very small metallic incrustations" in one core of pelagic clay[1].

Site 364 was drilled in a water depth of 2,449 m (see Fig. 1). The copper rests in moderate yellowish brown (10YR 5/4) to light olive grey (5Y 5/3) Miocene marly nannofossil ooze at a depth of ~ 152 m below the sea floor. Native copper was found in only the samples mentioned, although all other samples at the writer's disposal from this site (364) and the adjacent site (365) were carefully re-examined after copper was found. The sample at 22–24 cm contained five grains of copper whereas the sample at 19–21 cm contained several dozen. My data refer only to the latter sample.

Native copper makes up 12% (by weight) of the total sample. The copper grains are metallic reddish-brown, mostly botryoidal masses, and individually range up to 4 mm across. A number of copper casts of planktonic foraminifera (Fig. 2) are also present (the frontispiece of vol XL, *Initial Reports of the Deep Sea Drilling Project* carries colour photographs of several of these grains). Many of the grains of copper are partially coated with very light green calcium carbonate and clay; most, but not all, of this coating is soluble in dilute HCl. The greenish tint is probably caused by reaction of the copper with the surrounding carbonate-rich sediments, forming copper carbonate.

Table 1 shows that, with the exception of small amounts of sulphur, the grains can be considered wholly copper (the concentrations of the other elements, if real, are negligible). The presence of 0.18% and 0.64% sulphur suggests a very small amount of copper sulphide may be present. The copper value in the < 89-μm diameter size fraction is obviously vastly higher than in normal marine sediments[1], indicating that minute copper particles are also present in this fine fraction. Concentrations of the other elements in this fraction are more or less normal for marine sediments.

The presence of native copper in any natural environment (whether igneous, metamorphic, or sedimentary) is always somewhat surprising in view of the ease with which copper combines with other elements (especially sulphur). Nevertheless, copper abundantly exists in this form in a diversity of environments, and the literature on the subject is voluminous. Its presence in sedimentary rocks, the concern of this study, is less common than in igneous rocks, but the literature is still extensive[2,3].

Onshore erosion and mechanical transport of these particles to site 364 may be discarded as a possible mechanism of origin. It is difficult to imagine transport which would bring only copper grains to this site, without other terrigenous grains as well. Moreover, the copper grains are not rounded and abraded, as would be caused by any transporting mechanism capable of carrying these large grains. Finally, the presence of copper-infilled planktonic foraminifera conclusively proves an *in situ* deposition from copper solutions.
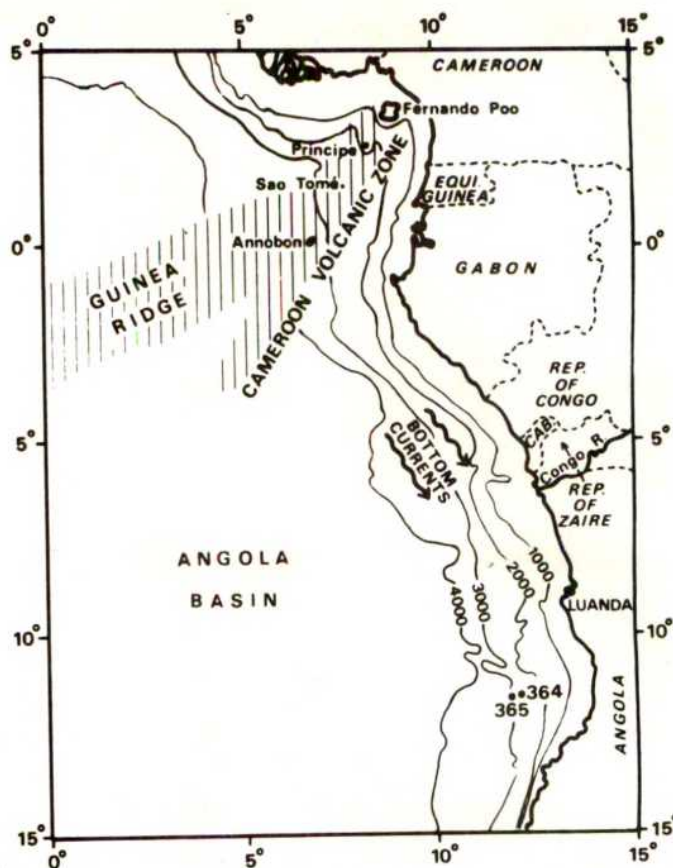
Fig. 1 Location map after Emery *et al.*[16] showing site 364, where native copper has been found, and site 365. Isobaths in metres.

Syngenetic origin of copper in marine sediments requires: first, a source of the metal; second, transporting solutions and a force to move them and third, a precipitating agent. Several conceivable origins for the metal: a deep-seated magma, with ascending hydrothermal solutions carrying copper, extraction and concentration of copper from seawater by microorganisms, whose death causes localised enrichment of copper in the sediments (for example, the Kupferschiefer Formation[4–6]), or, finally, copper solutions derived from a submarine basaltic extrusion and transported to the site by currents.

Fig. 2 Electron microprobe scan showing homogeneous distribution of copper in a planktonic foraminifer cast. All light areas are copper; dark areas within the cast are unfilled spaces.
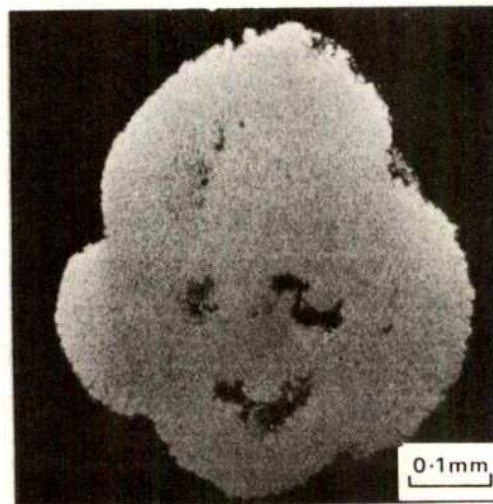
Several lines of sedimentological and chemical evidence militate against the first two possibilities', and the third possible source is the most attractive. This type of source has been described as "volcanic emanations" or "hydrothermal exhalations" by many authors. Geochemical evidence has been presented which would seem to prove the presence of submarine hydrothermal solutions⁸. The margins of hot, submarine-erupted basaltic rocks are quickly cooled, forming pillow structures whose chemistry approximates the chemistry of the erupted liquid. The slowly cooled interiors of the flows are, on the other hand, depleted in certain elements which are known to be enriched in pelagic sediments and manganese nodules⁸. This depletion is caused by seawater entering the eruptive flow through contraction fractures, being heated, and dissolving volatiles and metals concentrated in the residual phases within the interior of the extrusion. Metals are then mobilised as chloride complexes⁸. This is in accord with work showing that a Na–Ca–Cl brine is a potent solvent for copper and other metals⁸.

This heavy, metal-enriched brine could be expected to flow over the sea floor, directed by gravity, bottom currents, or both. It is tempting to consider the north-east–south-west trending Cameroon Volcanic Zone as the possible extrusive source. This volcanic zone extends almost 600 nautical miles off-shore, the terminal end being < 700 nautical miles north of site 364 (Fig. 1). Moreover, the offshore segment is known to have extruded basalt during the Miocene⁹. Present-day movement of bottom currents in the eastern Angola Basin is to the south¹⁰. If a similar circulation prevailed during the Miocene —and it is believed that Neogene current patterns were similar to those of today¹¹—transport of the brine to the site 364 area would have been the result.

No matter where a copper solution came from, or how it eventually reached site 364, some agency caused it to precipitate native copper. Organic matter has been suggested as a possible agent¹⁸,³. There have been several reports of native copper forming in modern peat swamps rich in organic matter¹¹⁻¹⁴. This is especially intriguing in view of the occurrence of authigenic marcasite at site 364 (ref. 15). Marcasite is a

Table 1 Analysis of selected elements

Electron microprobe (%)			X-ray fluorescence (p.p m)	
Foraminifera		Mass	< 89-μm fraction	
Specimen 1	Specimen 2			
Cu 99 92	99.08	99.83	Cu	3,029.38
S 0.18	—	0.64	Sr	556.20
Ca 0.08	—	0.01	Zn	204.75
Mg 0 05	—	0 01	Rb	95.46
Al 0 05	—	0.01	Ni	90.47
Fe 0.04	—	0.06	Zr	74.32
K 0.02	—	0.00	Y	20.62
Mn 0.01	—	0.01	Nb	6.92
Si 0.00	—	0.03		
Ti 0.00	—	0.01		
100.35		100 61		

mineral normally associated with peat swamps and is unusual in the marine environment. Its sporadic occurrence at site 364 suggests periodic conditions of extreme acidity. Bacteria may assist organic matter in precipitating native copper in acidic swamps. Indeed, metallic copper has been produced experimentally from copper solutions by using swamp bacteria¹⁸. Apparently native copper is formed by the reaction of bacterial waste products with copper solutions¹⁸ A concentration of organic matter may have been the precipitating agent at site 364, even though the sediments are not enriched in organic matter today.

Many questions still remain unanswered. Why, for example, are no other metals or their sulphides (other than pyrite) found at this horizon? Surely hot brine would have dissolved

other metals from basalt, if this were the source. Perhaps the fact that the common metals precipitate (as sulphides) in the order Cu, Zn, Pb, Ag may be relevant. Could copper have been precipitated here, early and alone, then the brine transported elsewhere to deposit other metals?

The foregoing discussion of the origin of this copper is highly speculative; it would be naive to think that a final and complete answer can be given yet. Nevertheless, it is hoped that the hypotheses presented will stimulate further ideas on the genesis of this unusual copper deposit.

I thank Mr R. S. Rickard for operating the microprobe and Professor R. V. Dingle and Mr J. H. W. Ward for helpful comments.

W. G. SIESSER

Department of Geology,
University of Cape Town,
Rondebosch, 7700, South Africa

1 Donnelly, T. W., and Nalli, G., in *Init. Rep. Deep Sea Drilling Project*, XV (US Government Printing Office, Washington, 1973)
2 Lindgren, W , *Mineral Deposits*, 930 (McGraw-Hill, New York, 1933)
3 White, D. E , *Econ Geol* , 63, 301–335 (1968).
4 Brongersma-Sanders, M., *Geol Rdsch* , 55, 365–375 (1965)
5 Brongersma-Sanders, M., *Miner. Deposita*, 3, 286–291 (1968)
6 Brongersma-Sanders, M , in *Sedimentary Ores Ancient and Modern*, 61–71 (University of Leicester Press, 1969).
7 Siesser, W. G , in *Init. Rep. Deep Sea Drilling Project*, XL (US Government Printing Office, Washington, in the press)
8 Corliss, J. B., *J geophys. Res* , 76, 8128–8138 (1971)
9 Baker, P E , in *The Ocean Basins and Margins The South Atlantic*, 493–533 (Plenum, New York, 1973).
10 Shannon, L V., *Fish. Bull S Afr* , 6, 27–33 (1970)
11 Hay, W. W , in *Studies in Paleo-Oceanography*, 1–5 (SEPM Spec Pub 20, Tulsa, 1974).
12 Lovering, T. S., *U S. geol. Surv. Bull* , 795, 45–52 (1927)
13 Forrester, J D , *Econ Geol* , 37, 126–135 (1942)
14 Eckel, E. B., *U.S geol. Surv. Prof Paper* 219 (1949)
15 Siesser, W. G , *Conf Marine and Freshwater Res S Afr.*, Fiche 6 E7-G2 (1976).
16 Emery, K O , Uchupi, E , Phillips, J , Bowin, C , and Mascle, J , *Woods Hole Oceanog. Inst Tech Rep. WHOI-74-99*, Unpub Ms , 152p (1974)

Experimental evidence that oxygen is the principal impurity in natural diamonds

IN virtually all treatments of the physical properties of natural diamonds, nitrogen is considered to be a major impurity and is used to explain many properties (see ref. 1). This is true because of the analysis of ref. 2 where a good correlation between nitrogen concentration and optical spectra was found. Those results were obtained from the conversion of diamonds to graphite in a carbon crucible. The report that nitrogen is the major impurity in type 1 diamonds can be tested indirectly and qualitatively by analysing occluded gases released by crushing such diamonds. Numerous experiments in this laboratory on diamonds from Africa, North and South America have failed to show that nitrogen is the major impurity in natural diamonds³,⁴ and we decided to reinvestigate the composition of gases released by the conversion to graphite of natural diamonds. We found oxygen and hydrogen to be the major impurities, and nitrogen to have only a minor role.

The diamonds used were both type I and type II (as characterised by ultraviolet absorption⁵), and free of detectable inclusions under binocular microscope examination at 90× Their body colour ranged from colourless to pale yellow and their weight ranged from 0.070 to 0.200 g. Possible contamination of the gases released from the surface of the apparatus was minimised by transforming the diamond crystals in a resistance-heated rhenium filament located near the centre of a large Nonex glass vacuum chamber. The vacuum chamber was attached to the inlet system of a research mass spectrometer⁷. Before being converted to graphite the diamond crystals were maintained at a temperature of 800 °C and a vacuum of 10⁻⁸ mmHg for several hours to remove possible surface contamination

Table 1 Analysis of gas released by the conversion under vacuum of a natural type I diamond (0.143 g) to graphite

Gas $T(°C)$	Percentage abundance (volume)						
	1,500	1,800	1,900	2,000	2,200	2,700	2,900
H_2	8.69	7.60	6.53	3.20	2.42	7.41	4.65
CH_4	0.46	0.22	0.27	0.16	0.26	0.27	0.14
H_2O	0.89	3.03	9.90	4.15	2.39	3.02	5.12
CO	78.98	87.12	79.57	90.67	90.88	84.68	86.09
N_2	1.05	0.86	0.90	0.78	0.82	0.73	0.79
Ar	0.03	0.02	0.02	0.02	0.02	0.01	0.02
CO_2	9.90	1.15	2.81	1.02	3.21	3.88	3.19
Volume* (cm^3 STP)	0.05	0.06	0.06	0.07	0.05	0.03	0.01

*These volumes correspond to the following average weight % of atoms: $H = 10.0\%$, $C = 28.1\%$, $N = 2.5\%$, $O = 59.2\%$ and $Ar = 0.2\%$.

The several natural diamonds, both type I and type II from Arkansas, USA used, gave relatively consistent results. Analytical data from a typical experiment on a type I diamond are given in Table 1. The gases released from the diamond, in decreasing order of abundance, were CO, H_2, H_2O, CO_2, N_2, CH_4 and Ar. The average atomic weight % is $O = 59.2$, $C = 28.1\%$, $H = 10\%$, $N = 2.5\%$ and $Ar = 0.2\%$. These data are consistent with the gases released by the crushing of numerous other diamonds *in vacuo*[3,4].

To determine the amount of contamination derived from the apparatus, two different approaches were used: first, control experiments with no diamonds present were carried out to determine the amount of gas released from the apparatus and second several diamond crystals were converted to graphite at lower temperatures using platinum–rhodium filaments. These experiments showed that the maximum level of contamination for nitrogen was $< 0.1\%$ and that for other gases $< 1\%$.

Since Kaiser and Bond[3] used graphite in their experiments, we conducted a control experiment with graphite. Gas release from a preconditioned sample of spectrographic grade of graphite at 2,000 °C showed nitrogen to be the major component.

This work was supported by the NSF.

CHARLES E. MELTON

Department of Chemistry,

A. A. GIARDINI

Department of Geology,
University of Georgia,
Athens, Georgia 30602

1 Berman, R., *Physical Properties of Diamond* (Oxford University Press, 1965).
2 Kaiser, W., and Bond, W. L., *Phys. Rev.*, 115, 857 (1959).
3 Melton, C. E., and Giardini, A. A., *Am. Miner.*, 59, 775–782 (1974).
4 Melton, C. E., and Giardini, A. A., *Am. Miner.*, 60, 413–417 (1975).
5 Clark, C. D., in *Physical Properties of Diamond*, 295–324 (Oxford University Press, 1965).

Nightglow and a new band system in molecular oxygen

I PROPOSE here that the transition $^3\Sigma^+_u \rightarrow {}^1\Delta_g$ in O_2 is responsible for previously unidentified features in the nightglow from the Earth's atmosphere, and in observations of radiation emitted from recombining atomic oxygen in the laboratory. The transition violates the approximate selection rules $\Delta S = 0$ and $\Delta\Lambda = 0$, ± 1 (ref. 1); it may be compared with the transition $^1\Delta_g \rightarrow {}^3\Sigma^-_g$, which, in addition to violating these selection rules, also violates the rigorous selection rule $g \leftarrow/\rightarrow g$ for electric dipole radiation, and so proceeds by magnetic dipole radiation[2]. The transition $^3\Sigma^+_u \rightarrow {}^1\Delta_g$ may therefore be significantly stronger than the transition $^1\Delta_g \rightarrow {}^3\Sigma^-_g$.

An accurate calculation of the wavelengths of the origins of the various bands of the $^3\Sigma^+_u \rightarrow {}^1\Delta_g$ system is not possible, since only the lowest two vibrational levels of the free $^1\Delta_g$ mole-

cule have been observed. In Table 1 we present a Deslandres array based on data given by Krupenie[3]. The entries underlined correspond to features which are clearly present in published nightglow spectra[4-7]. In particular the unexplained features pointed out by Chamberlain[4] and Hennes[6] are accounted for. The bands fall near a Condon locus which corresponds closely to that found for the Herzberg I system[8] ($^3\Sigma^+_u \rightarrow {}^3\Sigma^-_g$), where the Franck–Condon factors will be similar. Other bands of the $^3\Sigma^+_u \rightarrow {}^1\Delta_g$ system may well be present in the nightglow, but are unclear because they coincide with bands of the Herzberg I system, or because they are weak. Such possible identifications are underlined with dashed lines.

As in the case of the Herzberg I system, there are a number of coincidences and near coincidences in wavelength of band origins. This arises because the vibrational quanta of the $^1\Delta_g$ state are approximately twice as large as those of the $^3\Sigma^+_u$ state. In cases where two or more bands might be present on wavelength considerations alone, comparison with bands present in the Herzberg I system[8] provides a useful guide. Thus, for example, we identify the feature at 3,789 Å (refs 4 and 6) with the (3,2) band rather than the (1,1) band, although both are predicted to have an origin at 3,786 Å. Also, preferential population of the upper vibrational levels of $^3\Sigma^+_u$ is characteristic of the nightglow[9]. It is possible that emission from lower vibrational levels of $^3\Sigma^+_u$ to $^1\Delta_g$ is present in the nightglow, and contributes to the weakly banded blue pseudocontinuum[9], but identification is not possible.

Laboratory investigations of the radiation emitted by recombination of atomic oxygen[10] have shown the presence of numerous features between 3,500 Å and 5,000 Å which do not belong to the Herzberg I system. Degen[11] has presented a high resolution spectrum of a band with its origin near 4,007 Å, and stated that most of the other unidentified bands seem to belong to this system. The rotational structure is complicated,

Table 1 Calculated band origins in Å for the system $^3\Sigma^+_u \rightarrow {}^1\Delta_g$ in O_2

v'' / v'	0	1	2	3	4	5
0	3,687	3,900	4,135	4,395	4,685	5,008
1	3,584	3,786	4,007	4,251	4,521	4,822
2	3,491	3,682	3,891	4,120	4,374	4,655
3	3,407	3,588	3,786	4,003	4,242	4,506
4	3,330	3,503	3,692	3,898	4,124	4,373
5	3,261	3,427	3,607	3,803	4,018	4,254
6	3,199	3,359	3,531	3,719	3,925	4,150
7	3,144	3,298	3,464	3,645	3,842	4,057

Entries fully underlined seem definitely present in the nightglow; those partially underlined are probably present. Entries for $v' > 7$ are omitted because these levels of $^3\Sigma^+_u$ are believed to be absent in the upper atmosphere at night.

as is to be expected for the $^3\Sigma^+_u \rightarrow {}^1\Delta_g$ system, which, like the $^1\Delta_g \rightarrow {}^3\Sigma_g$ system, will have nine branches, three Q-form, two P-form, two R-form, one O-form, and one S-form (ref. 1). In general the calculated band origins in Table 1 are a few Å less than the band positions observed in the nightglow and in the laboratory. Thus we identify Degen's band at 4,007 Å with the (3,3) band. This may be partly because the bands are red degraded, and the levels corresponding to $J = 0$ and 1 in $^1\Delta_g$ are missing, and partly because the vibrational terms for $v > 1$ in $^1\Delta_g$ need revising. A successful analysis of the rotational structure observed by Degen in terms of the transition $^3\Sigma^+_u \rightarrow {}^1\Delta_g$ would not only confirm this identification but also give accurate values for the higher vibrational terms of $^1\Delta_g$. The hypothesis presented here seems superior to the proposal of Chamberlain[12] that some of these features are due to the $^3\Delta_u \rightarrow {}^1\Delta_g$ transition in O_2; not only do the wavelengths seem to agree better, but the $^3\Sigma^+_u$ state is already known to be present in laboratory systems and in the upper atmosphere at night. The contribution of either process to the population of the $^1\Delta_g$ state observed in the nightglow will be negligible.

I thank Dr M. Gadsden for comments.

P. C. WRAIGHT

Department of Natural Philosophy,
University of Aberdeen,
Aberdeen AB9 2UE, UK

1 Herzberg, G , *Spectra of Diatomic Molecules*, 240–242 (Van Nostrand, New York, 1950).
2 Van Vleck, J. H., *Astrophys. J* , 80, 161–170 (1934)
3 Krupenie, P. H., *J phys. Chem. Ref. Data*, 1, 423–534 (1972).
4 Chamberlain, J. W., *Astrophys. J* , 121, 277–286 (1955)
5 Krassovsky, V. I., Shefov, N. N., and Yarin, V. I., *Planet Space Sci*, 9, 883–915 (1962).
6 Hennes, J. P., *J. geophys. Res.*, 71, 763–770 (1966).
7 Broadfoot, A. L., and Kendall, K. R., *J. geophys. Res* , 73, 426–428 (1968).
8 Degen, V , and Nicholls, R W , *J Phys., B.*, 2, 1240–1250 (1969).
9 Chamberlain, J. W., *Physics of the Aurora and Airglow*, ch. 9 (Academic, New York 1961).
10 Brokla, H. P., and Gaydon, A. G , *Proc. R Soc* , A222, 181–195 (1954).
11 Degen, V , *Can. J. Phys.*, 46, 783–787 (1968).
12 Chamberlain, J. W., *Astrophys. J* , 128, 713–717 (1958).

Imagery, affective arousal and memory consolidation

A DISPARITY exists between the methods and theory that have been developed to study human memory and those used in animal learning studies. Present research strategies in human memory focus on (1) differences in the stimuli-to-be stored in memory (such as the extent to which words can be encoded)[1]; (2) conditions that precede new learning which influence the acquisition, stimulus transformation and encoding of input; (3) the context in which information is presented at the time of storage or retrieval from memory (where context determines the specific encoding strategy used to store and retrieve information)[2]; and (4) conditions that affect storage decay (such as those which might prevent rehearsal or precipitate "unlearning" of information stored in memory). The main theme of this research stresses

stimulus characteristics, often causally defined, as determinants of input–output relationships in memory. Few attempts have been made to develop the structural detail of human memory with respect to underlying biological mechanisms.

Studies of biological determinants and mechanisms of learning and memory have generally been on animals and have focused on neurochemical and neurophysiological conditions that alter learning or retrieval. Unlike human learning studies, these have involved physiological manipulations of brain after learning has taken place, at a time when experience is thought to be consolidated in memory. This research and its developing theory have emphasised the role of time- and intensity-dependent, specific and non-specific, brain state arousal in memory consolidation[3]. Unfortunately, few bridges exist between the methods and theory used in the biologically oriented animal learning studies and those thought useful in researching human learning. In this study we have attempted to integrate these two territories of findings and theory through manipulations of stimulus imagery and consequent encodability and of arousal using measures which in separate studies and approaches reliably affect recall probability.

Forty college students were first practised and then individually tested in a control condition and one of four experimental conditions. In the control conditon, subjects viewed eight different groups of four common English words, each group consisting of two easily and vividly imageable words and two words difficult to image[1]. After viewing each word group for 4 s, subjects were asked to produce a single representative word which would serve to integrate each cluster of four words and to rehearse, or think about, the to-be-remembered words, for 10 s. After the subjects had seen and rehearsed all of the 32 words, they were engaged in a non-verbal distracting task for 10 min and then attempted to recall the presented words.

Different groups of these same subjects, ten in each, were also tested under one of four experimental conditions, constituting an imagery mnemonic and/or affective arousal manipulation superimposed on the control procedure. In all of these experimental conditions, subjects were provided with pictures projected from slides which were designed to help them organise input and aid recall. Pictorial stimuli depicted common human interactive themes, and were standard but ambiguous stimuli with known response characteristics [Thematic Apperception Test (TAT) plates 1, 3BM, 4, 5, 6BM, 7, 11, 13MF] which are used clinically to elicit emotionally rich themes as a projective personality test. In two of the conditions, these stimuli were presented for 10 s before the 4-s presentation of each cluster of four to-be-remembered words. The two remaining conditions involved presenting word clusters first (4 s) followed by the 10-s exposure of these same TAT slides during the time that to-be-remembered words were being rehearsed in immediate memory. Finally, one of each of the above two groups viewed the TAT slides under an instructional set to "get

Table 1 Free recall of words as a function of the timing of imagery mnemonics and affective arousal experiment conditions

		T_1	T_2	T_1A	T_2A
Mean total word recall	HI	6.5±0.4	7.4±1.2	8.3±0.8	11.1±0.9
	LI	3.9±0.5	4.3±1.0	5.0±0.7	7.8±0.8
	W/C	1.8±0.1	2.2±0.1	2.2±0.1	2.6±0.2
Change in recall produced by experimental conditions compared with control performance	HI	1.6±0.7	0.5±0.9	1.0±0.8	3.7±1.3
	LI	1.3±0.6	1.0±0.8	1.2±1.0	3.8±0.9

Subjects were studied in the following conditions: a control involving word recall without the use of imagery mnemonics or affective arousal; four experimental conditions, two involving only the use of an imagery mnemonic before the presentation of to-be-remembered word clusters (T_1) or after the presentation of these four-word sets (T_2) and the remaining two conditions using the imagery mnemonic with the induction of affective involvement before the presentation of each word cluster, (T_1A) or after their presentation, during rehearsal (T_2A). HI, High imagery words; LI, low imagery words; W/C, average number of words per recalled cluster.

involved with the projected interpersonal theme, by thinking about your own personal, relevant experiences". This manipulated set had been pretested in an independent group of 20 subjects and was shown to produce significantly higher self-ratings along a continuous global scale[4] of affective involvement compared to a set of eight non-projective test stimuli. These subjective ratings were also significantly greater than those produced in response to a content-equivalent set of conditions in which no such instructional set was introduced. At the same time, these two sets of stimuli produced the same ratings on a global imagery scale used to measure the effectiveness of the stimuli as an imagery mnemonic.

The contrasting effects of these conditions on components of recall were examined by means of a series of ANOVA. Results are displayed in Table 1 which describes changes in: (1) recall of high compared with low imagery words; (2) total recall of words; (3) the number of words per cluster that could be recalled (a measure reflecting within group structure or salience where words per cluster measures different memory processes than are involved in accessing a higher order memory unit, cluster recall[2]). The number of clusters of presented information recalled, where at least one item from that cluster was remembered, is a measure that reflects retrieval of higher order memory units and access to these structures stored in memory[2]. The number of groups of words recalled is the quotient of total word recall and word per cluster recall.

Analysis of variance demonstrated the following patterns of statistically significant effects. Control conditions consistently produced lower total recall scores than the experimental conditions, an effect due to fewer clusters of words recalled rather than differences in words per recalled clusters ($F=9.62$; d.f.$=4$, 45; $P<0.01$). In all conditions, more high imagery than low imagery words were recalled ($F=7.63$; d.f.$=1$, 45; $P<0.01$). Apparently the use of the imagery mnemonic made both high and low imagery words more accessible in memory. Imagery seemed always to change the salience of the stored stimulus trace.

Affective arousal strongly facilitated recall when information was already in memory, during a period of rehearsal and consolidation ($F=8.14$, d.f. 3, 27; $P<0.01$). Affective arousal was not effective at the time of information storage itself. Post-storage arousal facilitated recall, but increased not the number of higher order memory units (clusters) which could be retrieved but the number of words recalled per cluster (intrastructure salience).

Generally, treatments used in typical human memory studies are effective in inducing changes in the recall of higher order memory units but have little effect on recall of elements within such structure. In spite of the changes in recall produced by imagery and post-storage arousal, retrieval strategies for accessing memory were not altered. For example, neither input serial position effects on recall were altered by any of the manipulations used here, nor was output sequencing or clustering, or the tendency to maintain the integrity of higher order memory units at time of retrieval, altered in control or experimental conditions.

Our findings on the effect of arousal on memory consolidation are consistent with those in animal studies in which arousal by foot shock or drugs after learning alters trace consolidation[3], and also with findings in some clinical studies where patients with disturbances in affect, for example, depression, show disturbances in the consolidation phase of memory which is reversible by drugs that alter affect and arousal (for example, L-dopa which influences the dopamine system in brain)[5,6]. These findings and the experimental design that was used to test effects of the timing of the imagery mnemonic in association with affective arousal also points out the value of bridging the methods and theories common to traditional human memory studies, with their emphasis on storage effects, and

the relatively greater focus on post-learning variables commonly used in biologically oriented animal learning studies.

HERBERT WEINGARTNER
BURL HALL
DENNIS L. MURPHY
WALTER WEINSTEIN

*National Institute of Mental Health,
Bethesda, Maryland 20014, and
University of Maryland,
Baltimore County, Maryland*

1 Paivio, A., *Imagery and Verbal Processes* (Holt, Rinehart, and Winston, New York, 1971).
2 Tulving, E., and Thomason, D. M., *Psych. Rev.*, **80**, 352–373 (1973).
3 McGough, J. L., and Herz, M. J., *Memory Consolidation* (Albion, San Francisco, 1972).
4 Eagle, M. N., and Mullikan, S., *Am. J. Psychol.*, **87**, 409–423 (1975).
5 Murphy, D. L., Weingartner, H., and Henry, G., *Psychopharmacologia*, **27**, 319–326 (1973).
6 Henry, G., Weingartner, H., and Murphy, D. L., *Am. J. Psychiat.*, **130**, 966–971 (1973).

Antagonism between visual channels for pattern and movement?

GEORGESON has reported[1] some ingenious psychophysical experiments designed to test the hypothesis that 'sustained' visual cortical cells sensitive to the orientation of contours are antagonistically coupled to 'transient' cells in the same cortical column sensitive to movement at right angles to those contours. After inspecting one or other, or both, of two counter-rotating dot patterns, his subjects reported seeing "complex patterns of radial lines or curves". These appeared either to rotate in the opposite sense to the stimulus, or to be stationary if both counter-rotating disks had been viewed simultaneously in superposition. Georgeson attributed their appearance to the rebound response of 'pattern channels' inhibited during stimulation by 'motion channels' of the kind presumed to be responsible for movement after effects (MAE). We have tested this conclusion by extending the range of stimulus velocities well above and below those used by Georgeson. Our results suggest that stimulation of motion channels is not necessary for the effect observed, so that its appearance in his tests does not confirm his hypothesis.

Our doubts arose because at the speed used by Georgeson (66 r.p.m.) the retinal image of his rotating dot patterns must have generated appreciable circumferential streaks; and these alone could be expected to give rise to a 'complementary image' of radial streamers[2,3,5] without invoking stimulus movement as such. This point was first recognised by Wilson[5], who in fact described the same phenomenon in the late 1950s, together with the method of using two counter-rotating textured fields as the inducing stimulus. He listed the following after effects as the speed of rotation of the superimposed fields was increased: "(1) A set of faint, fine concentric circles were observed. (2) These became fewer but stronger and broader. (3) They expanded even more, radial whiskers became visible on the outer circles, and a general impression of flicker was observed. (4) The circles were then seen to be entirely composed of fine radial 'whiskers' or 'sparks' and appeared to pulsate in diameter. (5) With further increase of speed the circles expanded (leaving) radii, or sometimes a lattice pattern of opposite spirals, filling the whole field."

Wilson noted that the induction of orthogonal after images by rapidly moving stimuli had been described in 1845 by Sir David Brewster[6], who observed it after staring from a moving railway carriage. (Wohlgemuth, in his classical monograph *On the aftereffect of seen movement*[4], tried to explain Brewster's statement that the after effect was "transverse" to the direction of the objective movement
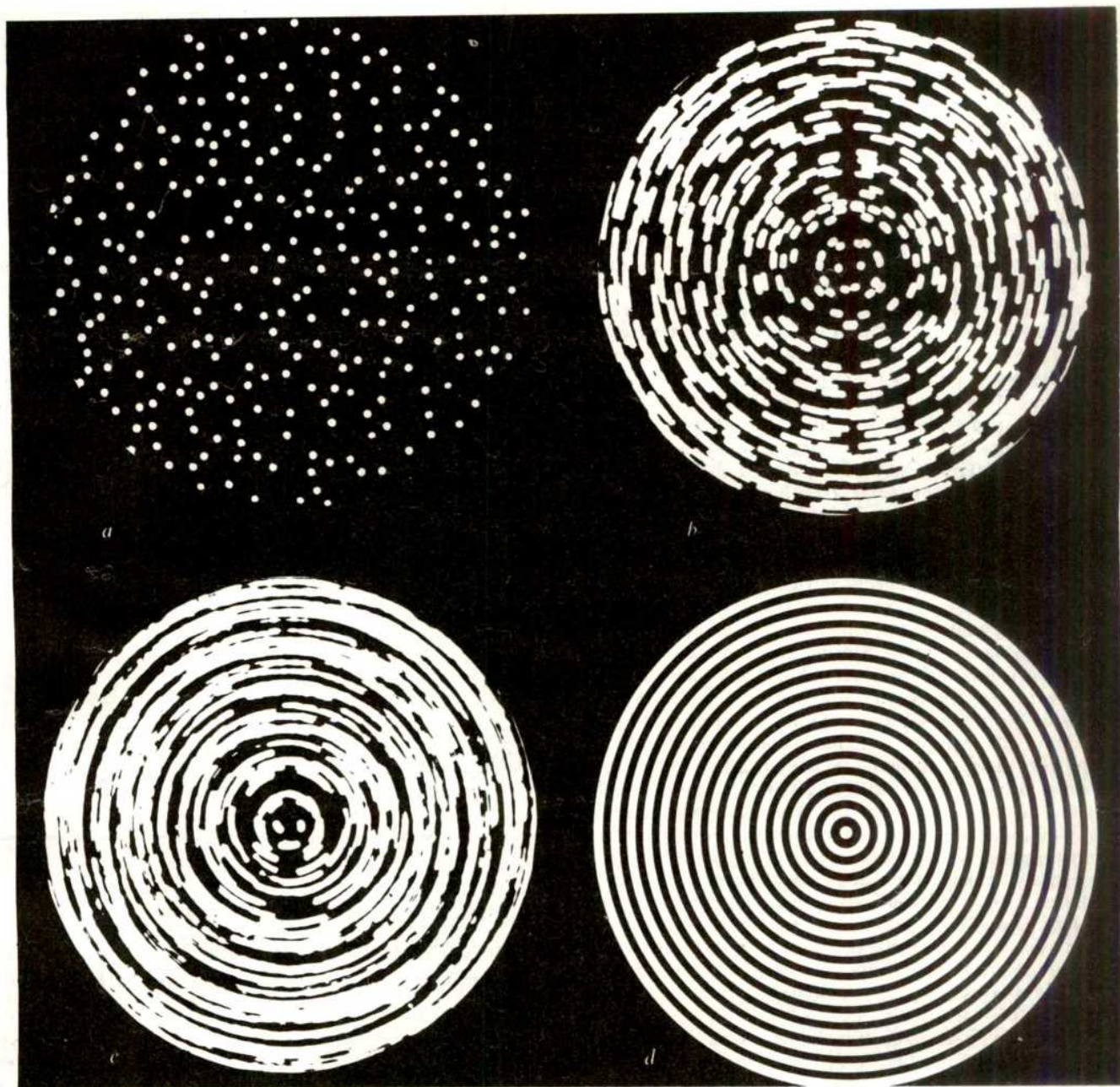
Fig. 1 *a*, Dot pattern used for counter-rotating stimuli; *b*, result of rotating *a* through 10° during exposure, and superposing mirror image; *c*, as *b* but 40° rotation; *d*, concentric-circle figure which excites a petalloid complementary image (see text).

as "possibly due to his changing the position of his head when he closed his eyes"!) Wilson later reported[7] that similar radial 'whiskers' can be induced by stroboscopic illumination of a single stationary circular contour.

Georgeson indicated that the 'movement analysers' he had in mind are those responsible for the familiar motion after effect or 'waterfall illusion'. Since the optimal rotation rate to generate a normal MAE is much lower than 66 r.p.m., human movement-sensitive channels should be more strongly stimulated if lower speeds were used. If Georgeson's interpretation were correct, this should intensify the radial lines seen as an after effect, and these might be expected to persist down to speeds at which the normal MAE disappears. Conversely, they should disappear at speeds above the upper threshold for perception of motion. If, however, radial lines were produced by the same mechanism as the normal complementary after image induced by inspection of stationary circular arcs (that is, by activation of contour-sensitive channels), they should persist at high rates of rotation, but disappear for rates that are too

low to produce significant circumferential streaks even though adequate for the normal MAE. We have therefore repeated Georgeson's observations, with eight subjects in all, using rates of stimulus rotation from 960 r.p.m. down to 3 r.p.m.

Our subjects viewed from 2.6 m a white screen on which two counter-rotating random-dot patterns were thrown from a 300-W projector having a motor-driven slide rotator. The two concentric images, produced with the aid of a Dove prism in front of the projector lens, were matched for brightness (110 cd m⁻² on white parts). They subtended an angle of approximately 20° at the subject's eyes, and were centrally fixated.

After an exposure of 1 min the subject immediately inspected a uniform white surface (70 cd m⁻²) and was asked to report any after images. For stimuli rotating at 60 r.p.m., all our subjects confirmed Georgeson's report of near-radial lines, usually adding that these seemed to "shimmer" or "stream" radially. Most subjects saw them extend almost to the centre, with only the central 4° clear of lines.

With a stimulus rotating at 30 r.p.m., however, six of seven subjects reported that the lines came only "about halfway in", leaving an 8° central area clear of radial lines. At 10 r.p.m. (which gave powerful MAEs with a single disk) only two of six subjects saw an after image of "faint radial whiskers" near the periphery of the stimulated field. At 6 r.p.m. all of seven subjects reported the whole test field free of radial lines, with only faint pulsating concentric rings as an after image, except that one subject (VM) detected faint radial "whiskers" around the periphery. Finally, three subjects tested at 3 r.p.m. could see no radial lines; yet the MAE at this rate (for a single disk) was still powerful over the whole area. As might be expected, the exact cutoff point for the radial after effect varies with individuals and with the rotating pattern used; but the trend of these results is clearly opposite to that predictable on Georgeson's interpretation.

Increasing the speed above Georgeson's rate of 66 r.p.m., we found no noticeable reduction in intensity of the chrysanthemum-like complementary after image up to 240 r.p.m. At 480 r.p.m. and 960 r.p.m., all observers agreed that it was considerably weaker, although still detectable; but at these speeds, the streaks produced by individual dots would be expected, and were observed, to run into each other and reduce effective contrast.

A rough quantitative check on this point can be gained by calculating the patterns of streaks to be expected if each retinal ganglion cell fired for 60 ms in response to a briefly flashed spot[8]. During this time at 60 r.p.m. each dot of a rotating pattern would trace an arc of about 20°. Figure 1a shows the dot pattern we used, and Fig. 1b and c the results of rotating it through 10° and 40° during photographic exposure and then superposing the mirror image, to simulate the streaks produced by counter-rotating fields at rates of 30 and 120 r.p.m. on the foregoing assumption. It is clear that (1) overlapping of streaks at these rates does not prevent a strongly circular pattern from developing; (2) at rates below 30 r.p.m. a progressively larger central area would become practically streak-free, and so ineffective in stimulating contour analysers; and (3) at rates above 120 r.p.m. (for our dot pattern) overlapping of streaks must progressively reduce effective contrast. This all fits quite well with the observations, and suggests that the assumed duration is of the right order of magnitude.

With a moving dot, the stimulating streak is effectively pulsed on and off as the dot passes over each receptive field, so the after effect would be expected to be (and is) stronger than that after inspecting Fig. 1b and c. Figure 1d (see ref. 2) has been included to let the reader verify that the complementary image induced by stationary concentric circles (or by a line of dots rotating at very high speed) is indeed an array of near-radial streamers. These may be seen either immediately on closing the eyes or on a uniform white surface. (At repetition rates of the order of 5–10 s⁻¹, radial streamers can be seen continuously during inspection. (Compare ref. 3, page 344.)

We conclude that Wilson's radial after effect is most parsimoniously attributed to stimulation of contour analysers by the rotating dots, and that stimulation of the motion analysers responsible for the MAE is not essential. This does not of course disprove the general idea of opponent coupling between 'sustained' and 'transient' cells. It was shown[2] that similar complementary images can be made visible continuously (not only as an after effect) by optically superimposing dynamic visual noise (for example, from a detuned TV receiver) on the fields of near-parallel contours, such at Fig. 1d[2,9]. The inference then suggested (before the discovery of contour-specific visual cells) was that channels sensitive to one orientation are rendered differentially hyperexcitable by 'satiation' of channels sensitive to the orthogonal orientation. Whether the mechanism was one of active mutual inhibition or of passive competition was left open by the psychophysical evidence; but in view of Wilson's finding[5] that the full effect requires the presence of about four to six near-parallel lines, and the fact that the wavelength of the sinuous illusory contours is typically about 4 times the inducing linespacing, it seems likely that network properties as well as single-unit characteristics are involved[2,3]. Georgeson's suggestion of antagonism between pattern- and motion-signalling channels is geared to the idea that removal of the parallel-line inducing pattern 'disinhibits' motion channels. The orthogonal patterns seen in dynamic noise suggest that even in the presence of the inducing lines, orthogonal channels for both pattern and motion may become hyperexcitable. In any case, we agree with Georgeson that it seems attractive to invoke excitation of motion channels in order to account for the streaming motion observed in the after effect, even if more complex relationships must be added to the picture to do justice to all the facts.

D. M. MacKay
Valerie MacKay

*Department of Communication,
University of Keele,
Staffordshire ST5 5BG, UK*

1 Georgeson, M. A., *Nature*, 259, 413–415 (1976).
2 MacKay, D. M., *Nature*, 180, 849–850 (1957); 181, 362–363 (1958).
3 MacKay, D. M., in *Sensory Communication* (edit. by Rosenblith, W. A.), 339–355 (MIT and Wiley, New York, 1961).
4 Wohlgemuth, A., *Br. J. Psychol., Mongr. Suppl.*, 1, No. 1 (1911).
5 Wilson, J. P., thesis, London Univ. (1960).
6 Brewster, Sir D., *Trans. Brit. Assoc.*, 8 (1845).
7 Wilson, J. P., *Nature*, 187, 137 (1960).
8 Levick, W. R., and Zacks, J. L., *J. Physiol., Lond.*, 206, 677 (1970).
9 MacKay, D. M., *J. gen. Psychol.*, 72, 181–197 (1965).

Evidence for a low upper limit of heritability of mental test performance in a national sample of twins

THE use of kinship correlations is a long established approach to the estimation of heritabilities (h^2) and component variances of human phenotypes. Among the best known studies are those using monozygotic and dizygotic twin correlations, data from which are commonly interpreted to represent a substantial heritability (that is, about 0.8) of IQ and other mental test performances (see Jensen[1] for review, and Jinks and Fulker[2] for the empirical requirements of this approach). Some doubt has been cast, however, over the empirical sufficiency of the most quoted estimates. Among the doubts are those concerning biographies of subjects, testing procedures, and problems emerging from the use of varied and inappropriate tests and from the subjects being of widely different ages[3]. More recent estimates using a version of the identical–fraternal twin comparisons (but derived from a sample in which the zygosities were unknown) have arrived at a value for h^2 not significantly different from zero[4–6]. The National Child Development Study (for details, see ref. 7), a longitudinal study of all the children in England, Scotland and Wales born in one week of March 1958, includes a nationally representative sample of twins about whom have been gathered considerable biographic, biometric, social and psychometric data (unpublished results). We have used this sample to investigate the broad heritability of performance on a general mental ability test.

Zygosities of twins were established during an interview of parents of twins by health visitors at the 7-yr follow-up and from responses to a letter sent to parents at the time of the

Table 1 Means (±s.d.) of verbal (V) and non-verbal (NV) test scores by zygosity and sex

| | MZ | | DZ$_{ss}$ | | DZ$_{os}$ | | Whole cohort | |
	V	NV	V	NV	V	NV	V	NV
Male	19.857±9.100	20.405±8.619	18.983±8.865	19.367±6.978	17.750±10.401	19.050±8.115	21.044±9.466	20.754±7.681
	n = 42		n = 60		n = 40		n = 7,262	
Female	18.825±8.743	19.400±6.719	22.300±8.339	21.360±7.016	20.175±7.576	19.725±7.791	23.124±9.128	21.015±7.536
	n = 40		n = 50		n = 40		n = 6,872	
t =	0.392 (NS)	0.484 (NS)	1.579 (NS)	1.174 (NS)	1.625 (NS)	0.526 (NS)	13.285; $P < 0.001$	2.038 $P < 0.05$

NS, Not significant.

11-yr follow-up; they were asked if the twins were 'identical' or 'non-identical' that is, 'fraternal'. This does not, of course, provide conclusive evidence of zygosity; however, several studies have shown that simple questionnaire responses of this sort reach at least 95% agreement with serological criteria, and it is generally felt that misclassifications by this method are uncommon[8-10]. The proportions of identical and fraternal twins classified in this way in the present sample approximate the theoretical values, and there were a few pairs, excluded from the present study, for whom the zygosities were unknown.

comparisons. The differences are 0.185 for the verbal test (that is, appreciably greater than MZ–DZss) and 0.117 for the non-verbal test. If we are correct in assuming only marginal sex bias in the test itself they probably reflect what is known as 'treatment effect', that is, inflation of the DZss correlation by virtue of parents, teachers, peers and so on treating same-sex twins in a similar way. Identical twins, as well as necessarily being of the same sex, also look alike and have other characteristics in common. Not surprisingly, it has repeatedly been demonstrated that identical twins are generally subject to more

Table 2 Analysis of variance tables for verbal and non-verbal test scores

| | MZ n = 41 pairs | | DZ$_{ss}$ n = 55 pairs | | DZ$_{os}$ n = 40 pairs | | DZ$_t$ n = 95 pairs | |
	V	NV	V	NV	V	NV	V	NV
MS within pairs	12.207	7.335	13.118	9.773	22.569	16.569	17.097	12.634
MS between pairs	69.446	54.296	64.019	39.552	63.255	47.999	65.705	49.573
Intraclass correlation	0.701	0.762	0.659	0.604	0.474	0.487	0.587	0.594

The data considered here relate to the second follow-up study when the subjects were 11 yr old. At that time all the subjects were administered a timed (30-min), group test of general mental ability by their teachers in schools, following strict but simple instructions concerning timing and test circumstances. The test, comprising 80 alternate verbal and non-verbal items, was specially designed for survey work with children of this age group by the National Foundation for Educational Research (for further details of reliabilities and validities, see ref. 11). Although it has not been cross validated with better known standardised tests, the two sets of items comprising it would seem to be very similar to types used in typical IQ tests (Fig. 1). It is as well to draw attention to this because reported h^2 estimates for performance on scholastic attainment tests are often appreciably lower than those for general ability tests, although they vary widely depending on the precise nature of the test[12]. Because of this we have calculated separate correlations for the verbal and non-verbal performances.

Table 1 shows the mean scores and standard deviations for the whole cohort and for twins separated by zygosity and sex. (Suffixes ss and os signify same sex and opposite sex, respectively.) Overall these figures suggest a marginally better performance for girls relative to boys, including those in the dizygotic (DZ) groups. This may be due to a sex bias in the test, the main consequence of which would be to reduce the DZos correlations slightly. Table 2 contains analysis of variance tables for verbal and non-verbal scores of twin pairs. A major question is the magnitude and statistical significance of difference between correlations. The most important ones from the point of view of heritability are (monozygotic) MZ–DZss. These are 0.042 for verbal scores and 0.158 for non-verbal scores, neither of which is statistically significant. Such non-significant differences are in themselves supportive evidence for very low or zero heritabilities.

The differences between fraternal twins of the same and opposite sex (DZss–DZos) are interesting because they suggest a possible source of error in identical–fraternal twin

uniform treatment than fraternals[5,6]. It is therefore possible to invoke treatment effects as the basis for inflated MZ twin correlations.

Table 2 indeed illustrates what happens when these 'treatment effects' are incorporated into the estimation of heritability (t indicates the total obtained from both os and ss twins). The difference MZ–DZt is non-significant for the verbal test, but although remaining small, becomes significant for the non-verbal test; using Z scores we have

$$Z_{MZ} = 1.000; \quad Z_{DZt} = 0.683$$

$$\frac{\Delta Z(MZ - DZt)}{\sigma(Z_{MZ} - Z_{DZt})} = \frac{0.317}{0.190} = 1.668 \, (P < 0.05)$$

On the face of it these data yield a heritability of 0.336±0.189

Fig. 1 Examples of (a) non-verbal and (b) verbal items used in the mental ability test.
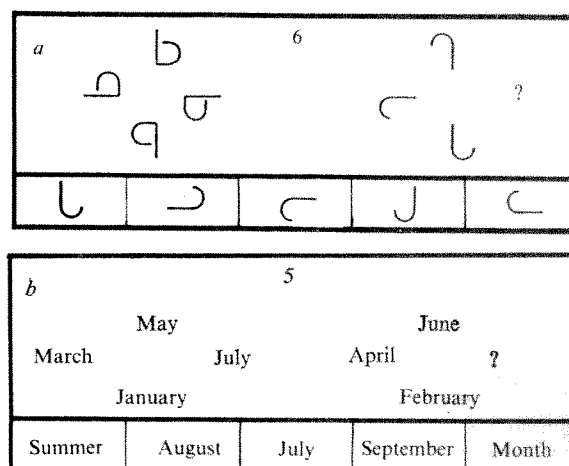
for non-verbal mental ability. This is obtained from the general expression, $h^2 = (r_{MZ} - r_{DZ})/(\rho_{MZ} - \rho_{DZ})$, where ρ_{MZ} and ρ_{DZ} are the theoretical correlations for identical and fraternal twins, respectively; essentially the derivation of Jensen[12]. Taking $\rho_{DZ} = 0.5$, of course, assumes no assortative mating (the presence of which would increase h^2) and no dominance or epistasis (which would decrease h^2). The magnitude of these influences is difficult to assess with accuracy from existing empirical data, but it is usually assumed that they are very small. Nonetheless, adopting Jensens'[12] estimate of $\rho_{DZ} = 0.55$ under assortative mating, and thus the maximum feasible correction to the above h^2, only increases it to 0.373. This h^2 is the consequence of introducing the usually unsuspected 'treatment effects' into the data and is at best an upper limit. It has been demonstrated that taking only minimal account of such effects in the data of previous studies reduces the reported h^2 estimates to values not significantly different from zero[5,6].

In conclusion, our nationally representative sample of twins of uniform age and test subjection, offer supportive evidence for zero or low upper limit heritabilities of mental test performance. Such estimates are considerably lower than the value of 0.8 normally quoted in psychometric research[13] and which has permeated the psychological literature. Although it is necessary to be clear about the often-inflated significance of heritability estimates[14], the present results are perhaps further indicative of the stringencies of sampling required in such studies[3].

We thank Professor Neville Butler for encouragement; Dr Joseph Schwartz, Dr Michael Schwartz and colleagues of the National Children's Bureau for comments; and the DES, DHSS and SSRC for support.

B. ADAMS
M. GHODSIAN
K. RICHARDSON

National Children's Bureau,
8 Wakley Street,
Islington, London EC1V 7QE, UK

1 Jensen, A. R., *Harv. Educ. Rev.*, 39, 1 (1969).
2 Jinks, J. L., and Fulker, D. W., *Psychol. Bull.*, 73, 311 (1970).
3 Kamin, L., *The Science and Politics of IQ* (Wiley, London, 1975).
4 Scarr-Salapatek, S., *Science*, 174, 1285 (1971).
5 Schwartz, M., and Schwartz, J., *Nature*, 248, 84 (1974).
6 Schwartz, M., and Schwartz, J., *Psychol. Bull.* (in the press).
7 Davie, R., *et al.*, *From Birth to Seven* (Longman, London, 1972).
8 Cederlof, R., *et al.*, *Acta Genet. Basel.*, 11, 338 (1961).
9 Cohen, D. J., *et al.*, *Arch. gen. Psych.*, 29, 465 (1973).
10 Bulmer, H. G., *The Biology of Twinning* (Clarendon, Oxford, 1970).
11 Douglas, J. W. B., *The Home and the School* (MacGibbon and Kee, London, 1964).
12 Jensen, A. R., *Proc. natn. Acad. Sci. U.S.A.*, 58, 149 (1967).
13 Eaves, L. J., and Jinks, J. L., *Nature*, 240, 84 (1972).
14 Lewontin, R. C., *Am. J. Hum. Genet.*, 26, 400 (1974).

Correction to Fisher's correlations between relatives and environmental effects

FISHER[1] has given a now much used model of assortative mating, and using formulae obtained for it he analysed data[2] on human height, span and forearm. In his model, phenotypic variance is regarded as the sum of three variances, additive, dominance and environmental; and genetic variance is the sum of the additive and dominance components. This was Fisher's first important contribution to genetics and it would seem that he did not appreciate fully the implications of his model. His formulae for parent–child correlation and sib correlation, in the presence of assortative mating, are not correct for his model. Moreover, the correct formulae for these correlations for his model suggest a different interpretation when the value for (genetic/phenotypic variance) is greater than 1. I present here a summary of criticisms of Fisher's formulae; details of which, with derivation of correct formulae will be presented elsewhere.

The additive variance is obtained by fitting a linear regression on the gene content of the genotype; dominance variance is the residual genetic variance, that is, genetic, additive variance. Fisher found that the value of his symbol c_1 defined as (genetic variance/phenotypic variance) was greater than 1 for height and span and concluded that this "gives no support to the supposition that there is any cause of variance in these growth features other than genetic differences". His method of partitioning the phenotypic variance into additive, dominance and other components has recently been used for IQ and estimates of its heritability (additive variance/phenotypic variance) have been the subject of some controversy.

Fisher realised that assortative mating will cause association between genotypes ("phases of factors" in Fisher's terminology). He devised a complicated method to take account of it and found that this association increases the genetic variance of the population. The increase in genetic variance is attributable entirely to an increase in the additive variance of the population[3].

Fisher did not consider what effect assortative mating has on the genotypes of a parent and his progeny. He assumed that the additive deviations are the only cause of resemblance between parent and child as he says, "Hence, since there is no association except of z (additive values) between parent and child, the parental correlation coefficient is

$$c_1 c_2 \frac{1 \times \mu}{2}$$

Here μ is the phenotypic correlation between husband and wife and c_2 is the total additive variance/genotypic variance; c_1 has already been defined.

Wright[4] on the other hand says that "Assortative mating introduces a correlation between dominance deviations of parents and offspring and between dominance deviations of either and additive deviations of the other". Fisher did not investigate such correlations and took no account of them. It can be shown that in Fisher's model of assortative mating there is a small correlation between the additive and dominance deviations of parent and child. Thus, the assumption on which Fisher obtained his formula for parent–child correlation is not correct.

To obtain his formula for sib correlation, Fisher discarded his model of assortative mating and reverted to random mating. He said, "The variance of a sibship, for example, depends, apart from environment, only on the number of factors in which the parents are heterozygous, and since the proportion of heterozygotes is only diminished by a quantity of the second order, the mean variance of the sibships must be taken for our purposes to have the value appropriate to random mating. . . .".

It is true that the variance of single sibship will be the same under random mating and assortative mating. The proportions of different types of matings will, however, be different under the two systems. As the mean sibship variance is Σ(frequency × sibship variance), it will not be the same under the two systems.

Moreover, Fisher adds the whole of the increase in variance, $A\tau^2/(1 - A)$ where A is the genetic correlation between husband and wife and τ^2 is the additive variance appropriate to random mating, to sib covariance and no part of the increase to sib variance. This is difficult to justify assuming Mendelian segregation on which his model is based.

Consider now his two formulae together

$$\text{Parent–child correlation} = \tfrac{1}{2} c_1 c_2 (1 + \mu)$$
$$\text{Sib correlation} = \tfrac{1}{4} c_1 (1 + c_2 + 2c_2 A)$$

In the presence of dominance, the sib correlation will be greater than the parent–child correlation, therefore

$$\tfrac{1}{4} c_1 (1 + c_2 + 2c_2 A) > \tfrac{1}{2} c_1 c_2 (1 + \mu)$$

which holds only if $(1 - c_2)/2c_2 \mu > 1 - c_1 c_2$. Fisher's model of

assortative mating will hold for those values of c_1, c_2 and μ for which this inequality holds. It can be verified that it holds only for certain combinations of values of these parameters. For example, if $\mu = 0.5, c_1 = 0.8, c_2 = 0.8$, we get $0.25 > 0.36$, which is not true.

These considerations lead to the conclusion that Fisher's formulae for parent–child and sib correlations are not correct for his model. It is, however, possible to obtain formulae using Fisher's model which do not suffer from these deficiencies. These formulae, in Fisher's notation, are

Parent–child
$$\text{correlation} = \tfrac{1}{2}c_1c_2[1 + A(1 - A)^2] + \tfrac{1}{2}c_1(1 - c_2)A(1 - A)$$
$$\text{Sib correlation} = \tfrac{1}{2}c_1c_2[1 + A(1 - A)^2] + \tfrac{1}{4}c_1(1 - c_2)$$

In both, the first term represents the contribution of additive deviations to correlation and the second term the contribution of dominance deviations. Note that the contribution of the dominance deviations to parent–child correlation is smaller than their contribution to the sib correlation since $A(1 - A) \leqslant \tfrac{1}{4}$. This is another reason why the concept of heritability defined as additive variance/phenotypic variance cannot be applied to human populations mating assortatively.

If the coefficient of assortative mating μ is known, the value of c_1 and c_2 can be obtained from the parent–child and the sib correlation formulae. Effectively, the difference between the two correlations gives a fraction of the dominance variance. The formulae given here and those obtained by Fisher, however, are for genetic correlations which are unknown and phenotypic correlations have been used instead. If the environments of sibs are more alike than those of parent and child, the difference between phenotypic correlations will also contain a fraction of the environmental variance. When the phenotypic correlational difference is too large to be explained by the amount of dominance variance in the population, the value of c_1 will come out to be greater than 1. Thus, such a value, far from being evidence for no environmental effects, indicates their presence. On the other hand, a value less than 1 for c_1 does not, of itself, exclude environmental effects, which could mean that the contributions of dominance deviations, as well as of environmental effects, are small.

The work reported here was carried out at Galton Laboratory, University College, London under the supervision of Professor C. A. B. Smith, to whom I am grateful.

Atam Vetta

Department of Science,
Oxford Polytechnic,
Oxford OX3 OBP, UK

1 Fisher, R. A., *Trans. R. Soc. Edinb.*, **52**, 399–433 (1918).
2 Pearson, K., and Lee, A., *Biometrika*, **2**, 357–462 (1903).
3 Vetta, A., *Ann. Hum. Genet.*, **39**, 447–453 (1976).
4 Wright, S., in *Quantitative Inheritance* (edit. by Reeve, E. C. R., and Waddington, C. H.), (HMSO, 1952).

Effects of artificial selection on reproductive fitness in *Drosophila*

A common feature of many selection experiments, when polygenically determined traits are involved, is a reduction in the reproductive fitness of the selected strains. Experimental selection for differences in the numbers of abdominal bristles and sternopleural chaetae in *Drosophila melanogaster* produced sterility and a reduction in fertility in the selected strains[1,2]. Latter[3] found that the "competitive index" (mating propensity, female fecundity and survival ability) had fallen sharply in strains of *D. melanogaster* subjected to experimental selection for differences in scutellar bristle number. Selection in *Tribolium* for changes

in developmental rates produced alterations in productivity, body size and viability of the flour beetles[4]. The reproductive fitness of chickens under selection for increased shank length was reduced steadily throughout the selection process[5].

Lerner[6] proposed that an increase in homozygosity is directly responsible for the reduction in reproductive fitness found in selected strains. He submitted evidence suggesting that heterozygosity is of prime importance in producing a balanced, highly fit natural population. This preselection phenotype is the outcome of numerous generations on which natural selection has operated to produce maximum reproductive fitness. As experimental selection proceeds, the balanced phenotype and genotype favoured by natural selection is disrupted. The selection process favours homozygosity for those genes which influence the selected trait. Homozygosity is also increased by incidental inbreeding, a consequence of the experimental selection procedure. On relaxation of the experimental selection pressures, natural selection is free to restore a balanced phenotype. The most heterozygous individuals are again favoured, causing the "gains" made in the experimental selection scheme to be lost. Lerner has called this phenomenon genetic homoeostasis and defined it as "the property of the population to equilibrate its genetic composition and to resist sudden changes".

Selection for divergent geotactic and phototactic maze behaviour in *Drosophila* is also thought to decrease reproductive fitness. The evidence is indirect and based on the observation that suspension of selection pressures in strains of *D. pseudoobscura* resulted in the reversion of the positive and negative geotactic and phototactic strains towards their original neutral maze response[7]. No decrease in productivity has, however, been found in divergent geotactic strains of *D. melanogaster* during 10 generations of selection[8]. This present study was undertaken to investigate directly the reproductive fitness of strains of *D. melanogaster* selected for positive and negative geotactic and phototactic maze behaviour.

Fifteen-unit classification mazes[9,10] were used to select experimentally for divergent geotactic and phototactic behaviour. Maze scores may vary from 1 to 16 on both geotactic and phototactic mazes. A score of 1 represents the most negative maze response possible and the maximum positive maze response results in a score of 16. A cage population of *D. melanogaster* which was established by hybridising large samples of 20 different wild-type strains, was used to begin selection for opposite geotactic and phototactic behaviour. Divergent behavioural selection was carried out by collecting approximately 300 females and 300 males which were run separately through either the geo-mazes or photo-mazes each generation. Sixty pairs of the most positive and negative individuals were selected as parents to initiate subsequent generations for both types of behaviour. Non-virgin females were run through the mazes so females selected as parents were deseminated by being placed at $-10\,°C$ for 10 min[11]. The geo-strains and photo-strains thus produced were highly divergent having undergone opposite selection pressures for over 40 generations. The degree of divergence is reflected in the mean geotactic and phototactic scores of the four strains. Following 42 generations of geotactic selection the mean score of the geonegative strain was 1.85 and the mean score of the geopositive strain was 14.90. The photonegative strain had a mean phototactic score of 2.01 while the mean score of the photopositive strain was 12.7. The unselected control strain was approximately neutral for both geotactic and phototactic behaviour throughout the selection process.

The following procedure was followed to estimate the egg-to-adult survival, egg hatchability, larva-to-adult survival and female fecundity of the strains under study. Virgin females and males were collected over a 48-h

Table 1 Egg-to-adult survival, egg hatchability and larva-to-adult survival of flies from the geotactic strains and phototactic strains

Strain	Mean % egg-to-adult survival	Mean % of eggs hatching	Mean % larva-to-adult survival (uncrowded)	Mean % larva-to-adult survival (crowded)
Geopositive				
Generation 42	56.6			
Generation 46				
Replication 1	57.6	81.9	91.4	70.0
Replication 2	58.0			
Geonegative				
Generation 42	74.2			
Generation 46				
Replication 1	68.9	92.7	88.8	76.4
Replication 2	69.3			
Photopositive				
Replication 1	61.6	87.6	92.2	74.1
Replication 2	68.2			
Photonegative				
Replication 1	65.2	86.8	88.7	78.6
Replication 2	71.2			
Control				
Tested at				
Generation 42	78.1			
Replication 1	72.8	91.7	96.3	83.5
Replication 2	78.7			

period and stored in separate half-pint food bottles. At the end of this period 25 pairs of adult flies (an average of 1 d old) were immobilised with ether and placed in an empty half-pint milk bottle. A food surface was placed on the open end of the milk bottle. Each food surface consisted of a Stender dish top filled with 5 ml of the standard medium covered with a circular piece of paper towelling coated with a thin suspension of bakers' yeast. The bottle was inverted after the flies' recovery from the ether treatment. At 12-h intervals thereafter, the bottle was turned right side up and the flies were gently knocked to the bottom of the bottle. The food surface was changed and the bottle was returned to the inverted position. The eggs that were laid during the preceding 12 h were then counted. The first 100 eggs counted on each surface for each 12-h period were transferred into a 37-ml shell food vial in groups of 10. This procedure was continued for 20 d as the females aged from 2 d to 21 d old. Approximately 13 d after each egg transfer the number of adults to eclose from each vial was recorded, giving a measure of egg-to-adult survival.

The number of eggs to hatch was estimated during the egg-to-adult survival experiment when 2 replications were tested for each strain. In this case, at 3-d intervals, the food surfaces were changed prematurely after 3 h. One hundred eggs from each strain were transferred on to a fresh food surface and the number of eggs to hatch was recorded 30 h later. A total of 700 eggs was transferred for each strain. The set of eggs produced during the following 9-h interval were used in the egg-to-adult survival experiment.

Larva-to-adult survival was estimated by transferring 50 first instar larvae recently hatched (from the above food surfaces) into food vials. The number of adults to eclose from these vials was recorded 13 d after the larvae transfer. A total of 350 larvae was transferred for each strain tested.

Initially, a single replication of the positive and negative geotactic strains (generation 42) was investigated for egg-to-adult survival and female fecundity. Subsequently, the egg-to-adult survival, number of eggs to hatch and larva-to-adult survival were measured for the phototactic, geotactic and control strains with 2 replications from each strain being tested for egg-to-adult survival. The egg-to-adult survival values (Table 1) were analysed by calculating a corrected sum of squares value for the replication

means within each strain. These values were summed over all strains and divided by 8, the total number of degrees of freedom, to approximate the variance of a replication mean. A t test (d.f.$=8$) was then used to compare the different paired replication means.

Table 1 reveals that egg-to-adult survival was significantly reduced in members of the geopositive strain compared with flies from the control strain ($t=5.70$, $P<0.01$) and the geonegative strain ($t=3.59$, $P<0.01$). The geonegative flies were not, however, different from the control flies ($t=2.11$, $P>0.05$). Individuals from both phototactic strains were significantly lower in survival value than the control strain flies (photopositive, $t=3.44$, $P<0.01$; photonegative, $t=2.40$, $P<0.05$) but they were not significantly different from each other for this fitness parameter ($t=1.05$, $0.5<P<0.3$).

The percentage of hatchings associated with the geopositive strain was also decreased by comparison with all other strains tested (Table 1). Little difference in larva-to-adult survival was found in the four experimental strains and the control strain. This suggests the decrease in egg-to-adult survival found in geopositive flies is likely due to decreased hatching success of their eggs. The egg-to-adult survival values are, however, on the average, only 0.8 as large as the products of the percentage of hatchings and larva-to-adult survival values. This probably stems from differences in the techniques involved in estimating these fitness parameters. For example, the larval density was greater in the egg-to-adult survival study which may have reduced larval viability and accentuated the between-strain differences, whereas in the larval transfer study the lower larval density resulted in more favourable conditions and larval viability may have been only minimally affected. In fact, by dividing the egg-to-adult survival values by the mean percentage of eggs hatching one can compute the average larval viability under crowded conditions (last column Table 1). This shows that the larval viability is reduced in all strains under the crowded conditions with the geopositive flies showing the lowest larval variability.

No differences were found in female fecundity among flies from the geotactic strains (generation 42) and the control strain. The mean number of eggs laid per female per d was 38.4 for the geopositive females, 34.6 for the geonegative females and 35.3 for the control females.

In summary, selection for divergent geotactic maze behaviour has reduced the egg-to-adult survival rate of the geopositive flies with no similar change observed in flies from the geonegative strain. This decrease in viability seemed to be the result of lowered egg hatching success and a decrease in larval viability in crowded cultures. A slight decrease was observed in egg-to-adult survival of both phototactic strains, but the photopositive flies were not different from the photonegative flies.

A decrease in reproductive fitness, as found in the phototactic flies, is not unexpected after experimental selection. The reduced egg-to-adult survival found in geopositive flies but not in geonegative flies is, however, surprising. What are the possible explanations for this differential decrease in reproductive fitness?

Latter and Robertson[12] also found an asymmetrical decrease in fitness in strains of *D. melanogaster* subjected to independent selection for 3 traits (abdominal bristle number, wing length and sternopleural bristle number). They observed that flies from the low selected strains were more reduced in reproductive fitness than flies from the high selected strains. It was concluded that some loci acted pleiotropically, influencing reproductive fitness and the selected phenotype. A reduction in fertility was noted by Wigan and Mather[2] and Mather and Harrison[1] while selecting for chaeta number in strains of *D. melanogaster*. Decreases in fitness were found to accompany the initial generations of selection, but after selection pressures were

relaxed, the fitness of individuals from the selected strains was observed to increase. These results led to the hypothesis that genes influencing fitness were linked with those being directly selected. Selection favouring unique gene arrangements that control the morphological trait being selected was felt to disrupt concomitantly the fitness controlling genes. The increase in fitness after suspended selection pressures may represent crossover events which "release" the polygenes under direct selection from the fitness controlling genes.

In my own experiment an additional possible explanation is that the different selective pressures associated with geopositive and geonegative selection may be responsible for the asymmetrical decrease in fitness. The lack of egg-to-adult survival rate differences between the photonegative and photopositive flies supports this idea. The phototactic maze is horizontal and members of divergent phototactic strains would traverse the maze expending equivalent amounts of energy. Because geonegative selection favours flies that can walk through the entire vertical maze opposing the pull of gravity, those flies must also be the hardiest and the most vigorous. Flies selected in the geopositive direction, in contrast, may include individuals that weakly stumble downwards, as well as those that walk purposefully down. Selection of the most energetic individuals in the geonegative direction may counteract the normally deleterious effects of experimental selection. While selection in both directions favours those individuals homozygous for geobehavioural genes, the concomitant selection for vigour in the negative direction may reduce the deleterious effects of inbreeding. King[13] studied inbreeding in rats and demonstrated that selection for vigour does offset its fitness-reducing effects. Selection of weak and stumbling flies in the positive direction may have directly favoured the least fit and most homozygous individuals.

DONALD W. PYLE

Department of Zoology,
Indiana University,
Bloomington, Indiana 47401

1 Mather, K., and Harrison, B. J., *Heredity*, 3, 1–52, and 131–162 (1949).
2 Wigan, L. G., and Mather, K., *Ann. Eugen.*, 11, 354–364 (1942).
3 Latter, B. D. H., *Genet. Res.*, 8, 205–218 (1966).
4 Dawson, P. S., *Genetics*, 37, 63–77 (1966).
5 Lerner, I. M., and Dempster, E. R., *Heredity*, 5, 75–94 (1951).
6 Lerner, I. M., *Genetic Homeostasis* (Wiley, New York, 1954).
7 Dobzhansky, T., and Spassky, B., *Proc. natn. Acad. Sci. U.S.A.*, 48, 1704–1712 (1962); 62, 75–80 (1969).
8 Watanabe, T. K., and Anderson, W. W., *Behav. Genet.*, 6, 71–86 (1976).
9 Hirsch, J., *J. comp. Physiol. Psychol.*, 52, 304–308 (1959).
10 Hadler, N. M., *Biol. Bull.*, 126, 264–273 (1964).
11 Novitski, E., and Rush, G., *Drosophila Inf. Serv.*, 22: 75–76 (1948).
12 Latter, B. D. H., and Robertson, A., *Genet. Res.*, 3, 110–138 (1962).
13 King, H. D., *J. exp. Zool.*, 26, 335–378 (1918).

A simple mechanism for population cycles

ALTHOUGH most animal populations exhibit irregular fluctuations, there has long been considerable interest in the dynamics of populations which seem to fluctuate in a cyclic manner with a period different from that of obvious environmental 'drivers'. We have analysed possible models of the mechanisms underlying the regulation of populations in order to simplify the extraction of useful information from specific cases of population oscillation.

Existing models of cycling populations involve one of three mechanisms:

● The population is assumed to be a component of a stable but underdamped system—that is, in the absence of environmental noise the system approaches equilibrium in a series of damped oscillations. Cyclic behaviour is interpreted as the response of such a system to random noise. The amplitude of the oscillations is then a random variable, the mean value of which depends on both the internal damping and the external noise. The period of the oscillations is determined largely by interaction within the system although the noise causes a steady loss of phase information. A mechanism of this type is implied in Moran's[1] representation of the Canadian lynx cycle.
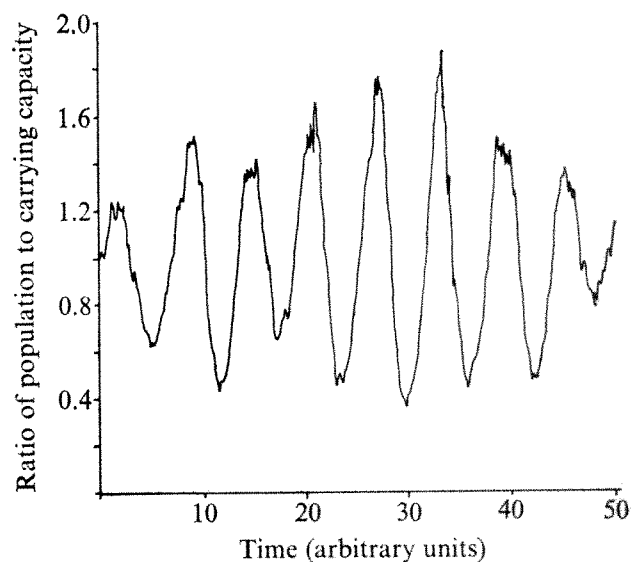
Fig. 1 Quasi-cyclic fluctuations in the time-delayed logistic model. The population is responding to fluctuation in the carrying capacity, K, of the form $K(t) = K_0(1 + \gamma(t))$ where $\gamma(t)$ is Gaussian white noise with standard deviation $\sigma = 0.1$. The intrinsic growth rate r and the time delay τ are chosen so that $r\tau = 1.4$, a value for which the system is underdamped.

● The population is again represented as part of a stable, underdamped system but the control mechanism is assumed to be nonlinear. In a strongly seasonal environment population cycles may then occur as a result of subharmonic resonance. We have previously[2] illustrated this effect in the time-delayed logistic model with strong periodic variation in the carrying capacity and Bigger (private communication) has suggested that the oscillations[3] in populations of the leaf miner *Leucoptera caffeina* (Washb.) may be of this type.

● The population is assumed to be limit cycling (possibly with some environmental fluctuations superimposed on the limit cycle). Both the amplitude and the period of the oscillations are then determined by the interactions within the system. Many examples have been discussed by May[4].

We have deliberately excluded from the above classification models like Bulmer's[5] of the Canadian lynx cycle which explain cycles in one population (lynx) as being driven by cycles in another population (snow-shoe hare) but which beg the question of the origin of these driving cycles.

The last two mechanisms—subharmonic resonance and limit cycling—involve complicated nonlinear effects and yet they have received considerably more attention from modellers than the first. Our aim here is to advocate the first mechanism as a paradigm for population cycles, since we feel that its obvious simplicity makes it a strong candidate for representing a seemingly cycling population in a fluctuating environment. Particularly appealing is the fact that the principle of the mechanism follows from a purely linear analysis.
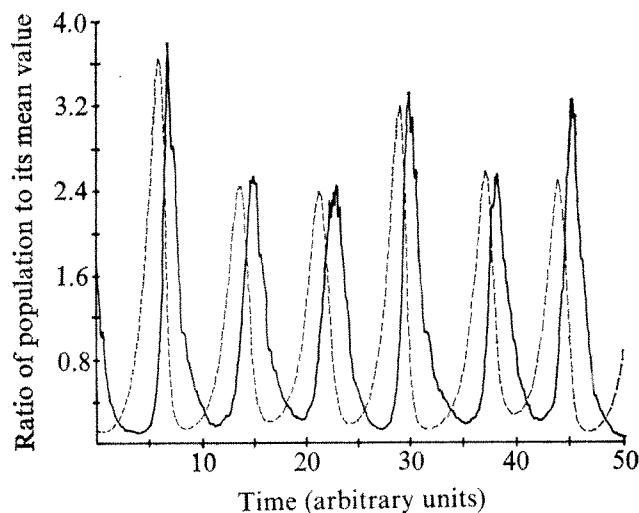
Fig. 2 Quasi-cyclic fluctuations in a damped Lotka–Volterra predator–prey model with a fluctuating predator death rate. The parameters are chosen to make the system underdamped. The continuous and dotted curves denote respectively the predator and prey.

We represent by $x(t)$ the difference between the instantaneous population value and its mean value, and by $\bar{x}(f)$ the Fourier transform of $x(t)$. The square of the modulus of $\bar{x}(f)$ is called the spectral density and is denoted by $S(f)$. If $p(t)$ denotes the fluctuation in some environmental parameter then provided it is small, the nature of the resulting population fluctuations can be deduced from the relationship

$$\bar{x}(f) = T(f)\bar{p}(f)$$

where $T(f)$ is known as the transfer function. We have previously[2] derived a transfer function for a simple population model and discussed its interpretation in some detail; here we need only note that an underdamped system is characterised by a peak in the transfer function. At high frequencies the transfer function of a stable system is always very small, essentially because a population does not respond to frequencies much greater than the reciprocal of the smallest ecological time constant in the system. Thus if (and only if) Fourier analysis of time series of relevant environmental variables reveals fluctuations which have a flattish spectrum over the range of frequencies for which the magnitude of the transfer function is significant, it is safe for the purpose of population analysis to represent these fluctuations as white noise. Where this idealisation is permissible, the population spectral density is proportional to $|T(f)|^2$.

Whenever environmental fluctuations lead to a peak in the population spectral density, quasi-cyclic behaviour results. Moreover, the argument in the previous paragraph guarantees that any underdamped system will respond in a quasi-cyclic manner to environmental fluctuations with a flattish spectrum. In Fig. 1 we show an example of quasi-cyclic population fluctuations obtained from a time-delayed logistic model with a time delay chosen to correspond to very light damping, and a randomly fluctuating carrying capacity. In Fig. 2 we illustrate the same behaviour in a damped Lotka–Volterra predator–prey model[6] with random fluctuations in the predator death rate. We have obtained similar cycles when analysing the effects of fluctuating mortality in a single-species model with age structure, and in a model of experiments[7] in which persistence in a spatially heterogeneous, predator–prey system is achieved through a balance of migration and local extinction. This work will be reported in detail elsewhere.

We have given a simple linear analysis of a mechanism which can generate oscillations. In principle, this treatment ought to be applicable only to very small oscillations, but we have carried out extensive numerical simulations of the above models and conclude that the linear analysis is remarkably accurate in predicting the size and nature of the population fluctuations right up to the point where extinction of the population on an ecological time scale becomes highly probable. Thus we conclude that population cycles of rather large amplitude may occur when an underdamped control system is subject to random environmental noise, and we suggest that this mechanism is worthy of consideration when analysing real population data.

We thank Mr M. Pettipher for assistance with the numerical analysis.

R. M. Nisbet
W. S. C. Gurney

*Department of Applied Physics,
University of Strathclyde,
Glasgow G4 0NG, UK*

1 Moran, P. A. P., *Aust. J. Zool.*, 1, 163–173 (1953).
2 Nisbet, R. M., and Gurney, W. S. C., *J. theor. Biol.*, 56, 459–475 (1976).
3 Bigger, M., *J. Anim. Ecol.*, 42, 417–434 (1973).
4 May, R. M., *Stability and Complexity in Model Ecosystems* (Princeton University Press, Princeton, 1973).
5 Bulmer, M. G., *J. Anim. Ecol.*, 43, 701–718 (1974).
6 Maynard Smith, J., *Models in Ecology*, 19 (Cambridge University Press, Cambridge, 1974).
7 Huffaker, C. B., *Hilgardia*, 27, 343–383 (1958).

Monogamy and duetting in an Old World monkey

THE first monogamous mating system among Old World monkeys has been discovered in the Mentawai langur, *Presbytis potenziani*. The only known monogamous anthropoid primates were the Old World gibbons (*Hylobates* spp.)[1] and a few New World monkeys[2-4]. Reproductive groups of all known Old World monkeys, represented by approximately 85 species, are either polygamous or promiscuous[5].

The endemic Mentawai langur is an arboreal folivore confined to Siberut, Sipora, North Pagai and South Pagai islands, 85–145 km west of Sumatra[6]. Our main study population was nine langur groups inhabiting a study area covering 200 ha in central Siberut Island, near the headwaters of the Sirimuri River (1°24′S, 99°1′E). Between July 1972 and November 1974 we spent 447 d and made over 160 h of visual observation of undisturbed Mentawai langurs at this site.

Adult Mentawai langurs exhibit little sexual dimorphism. Ten specimens (6 males, 4 females) show a female–male weight ratio of 0.98 (mean weight: male 6.5 kg; female, 6.4 kg). The only difference in coat coloration between the sexes is the male's conspicuous white genital fur patch.

The 21 groups of langurs we observed were adult pairs with up to four younger animals. Females bear a single infant approximately once every other year. Mean size of nine families studied regularly varied from 3.4 in March, 1973 to 4.0 following the 1974 July–August birth season. Adult pairs remained together for the entire 2.5-y study.

We did encounter six solitary males. These solitary males account for 22% (6/27) of our adult male sample. All adult females observed were monogamously mated. But all-male groups, typical of other polygynous langurs[4], were never encountered.

Families occupy a home range of about 15–22 ha. Home ranges of adjacent groups overlap extensively, but each contains an exclusively used "core area"[7] where their sleeping trees (mean = 2 per group) are located. Core areas are situated on major ridge crests, on boundaries between gibbon (*Hylobates klossii*) territories. Intrusion by alien langur males into core
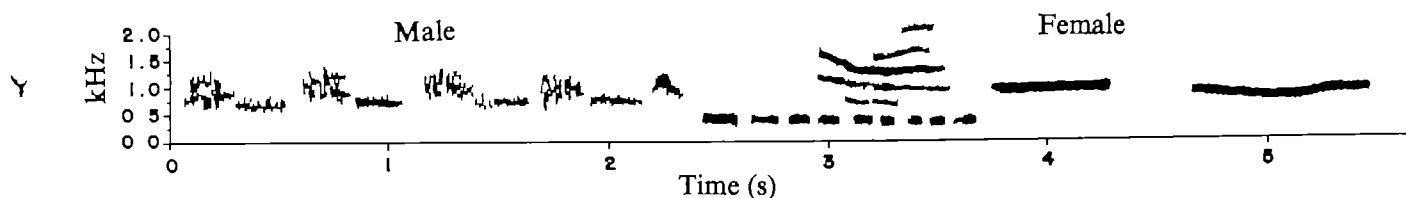
Fig. 1 Ink tracings of audiospectrograms illustrating a typical, duetted intergroup call by a mated pair of Mentawai langurs. The three prolonged (about 0 5 s) sounds terminating the duet are those of the female. The first sound overlaps the stacatto terminal trill of the male's vocalisation. Note the two components to the male call; a series of harsh, biphasic sounds corresponding to inhalation and exhalation, followed by a slow, tonal trill. The vocalisations were recorded in the field at 9.5 cm s with a Uher 4400 tape recorder and a Sennheiser 405-S microphone fitted to a parabola 60 cm in diameter. Audiospectrograms were made with a Kay Elemetrics sound spectrograph on wide band setting.

areas results in aggressive inter-male confrontations dominated by loud vocalisations and visual display by owner and trespasser.

Like adult females in other monogamous primates, female Mentawai langurs participate with their mates in vocal and visual displays directed towards adjacent groups. Males emit loud vocalisations (audible to about 1 km), either spontaneously or on hearing or seeing members of other groups. They are comparable to the intergroup "spacing" calls of other male colobines[5]. Male intergroup calls are accompanied by a vigorous jumping display on branches. This shakes the entire tree crown. When the males make intergroup calls, their mates respond by adding a characteristic 3–4-syllable coda (audible to about 0.5 km) to the end of it (Fig. 1). She then leaps from branch to branch, shaking the tree after the male has stopped his display. In other colobines such displays are given only by males[6].

The vocal portion of the Mentawai langur intergroup display, with its male and female components, resembles the duetting gibbon pairs[5], titi monkeys[2], and monogamous tropical birds which remain paired throughout the year[10,11]. Thus the correlation of duetting with permanent, year-round monogamy

Table 1 Behavioural response of Mentawai langurs to man

	Mated pairs with immature young	Mated pairs only	Solitary males
Distraction display	12	0	0
Cough alert	5	7	0
Rapid flight with no vocalisations	1	10	18
Total (n = 53)	18	17	18

in birds and primates can be extended to the Mentawai langur, and the hypothesis that duetting functions to maintain monogamous pair bonds[10–13] is supported.

Adult male Mentawai langurs perform distraction displays in response to man, the major and perhaps only predator on primates in Siberut. Other relevant predators such as large raptors or felids do not occur on these islands[6]. Another basis for the relatively great impact of man as a predator on Mentawai primates is the dominance of primates in the island's depauperate mammalian fauna Primates constitute 82% of all large mammalian prey taken by aboriginal hunters in Siberut. Thus, both the more intensive human predation and the lack of counterselection favouring alternative predator defence strategies probably account for the Mentawai langur's distraction performance in encounters with man.

The adult male distraction display is a loud vocalisation accompanied by branch bouncing, delivered every 20–30 m as he runs through the canopy. Vocalisations resemble the harsh, biphasic syllables of the intergroup call (Fig. 1), but the female remains silent. During this performance the female and immature young often hide silently and motionless in the canopy for periods up to 45 min before leaving. On five different occasions a displaying male made a complete circle after calling 6–8 times, entered the tree his family occupied, then again began a new distraction directed towards the observer.

The presence of offspring is significantly correlated ($\chi^2 =$ 51.9; $P < 0.001$) with the male's distraction display. For the 35 cases involving either solitary males or mated males with no offspring, langurs fled without giving alarm calls. Occasionally the male coughs harshly (7 of 17 events) if his mate is feeding in another tree. The cough (audible to about 50 m), may alert the female of danger and the male's intention to begin rapid movement. It may precede the louder call (audible to about 1 km) family males emit in response to man (Table 1).

Thus, several interrelated characters closely linked to monogamy in other species typify the Mentawai langur. Its lack of sexual dimorphism extends the correlation between monogamy and monomorphism[13,14]. The monogamous pair bond is maintained by vocal duetting and mutual visual displays The distraction display of family males is risky but increases the survival chances of his own offspring.

RONALD L. TILSON*

Institute of Ecology,
University of California,
Davis, California 95616

RICHARD R. TENAZA

Department of Biological Sciences,
University of the Pacific,
Stockton, California 95211

1 Tenaza, R. R., Folia Primat., 24, 60–80 (1975)
2 Mason, W A , in Primates, Studies in Adaptation and Variability (edit by Jay, P C), 200–216 (Holt, Rinehart and Winston, New York, 1968)
3 Moynihan, M , Smithsonian. misc coll , 146(5), 1–84 (1968)
4 Eisenberg, J F , Muckenhirn, N. A , and Rudran, R., Science, 176, 863–874 (1972).
5 Jolly, A , The Evolution of Primate Behaviour (Macmillan, New York, 1972).
6 Chasen, F N , and Kloss, C. B , Proc zool. Soc Lond , 0, 797–840 (1927)
7 Kaufmann, J H., Univ Calif Publs Zool., 60, 95–222 (1962)
8 Poirier, F. E , in Primate Agression, Territoriality and Xenophobia (edit by R L Holloway), 123–157 (Academic, New York, 1974)
9 Marler, P , and Tenaza R. R , in How Animals Communicate (edit by Sebeok, T A), (Indiana University Press, Bloomington, in the press).
10 Hooker, T , and Hooker, B I , in Bird Vocalizations (edit. by Hinde, R A), 185–205 (Cambridge University Press, Cambridge, 1969).
11 Thorpe, W H , Behav Suppl., 18, 1–197 (1972)
12 Lamprecht, J , Z Tierpsychol , 27, 186–204 (1970)
13 Wickler, W., The Sexual Code (Doubleday, Garden City, 1972)
14 Brown, J., The Evolution of Behavior (Norton, New York, 1975).

Sex chromosome translocations and speciation

SPECIATION is the primary process of cladogenetic evolution. It is generally agreed that genetic divergence sufficient to produce reproductive isolation among populations is

Table 1 Single-factor genetic basis of X·Y^L/ /Y^S male mating advantage

(A) Comparisons of yellow and normal bodied males

	X chromosome	Y chromosome	y^+mating frequency	n	χ^2 probability
(1)	X·Y^L, y^+ and X·Y^L	Y^S	0.53	45	0.655
(2)	X·Y^L	Y^S, y^+ and Y^S	0.50	260	0.902
(3)	X·Y^L, y^+f^+ and X·Y^L, f^+	Y	0.53	45	0.655
(4)	X, y^+f^+car$^+$ and X, f^+ car$^+$	Y^S	0.54	117	0.405
(5)	X, y^+f^+car$^+$ and X, f^+ car$^+$	Y	0.65	85	0.009*
(6)	X	Y^S, y^+ and Y^S	0.74	39	0.004†

(B) Comparisons of X·Y^L and X males

	X chromosome	Y chromosome	X·Y^L mating frequency	n	χ^2 probability
(1)	X·Y^L and X, f^+car$^+$	Y^S	0.57	44	0.366
(2)	X·Y^L, y^+ and X, y^+f^+car$^+$	Y^S	0.50	62	1.000
(3)	X·Y^L and X, f^+car$^+$	Y	0.40	50	0.157
(4)	X·Y^L, y^+ and X, y^+f^+car$^+$	Y	0.56	89	0.244

(C) Comparisons of Y^S and Y males

	X chromosome	Y chromosome	Y^S mating frequency	n	χ^2 probability
(1)	X·Y^L	Y^S and Y	0.63	143	0.003†
(2)	X·Y^L, y^+	Y^S and Y	0.69	45	0.017‡
(3)	X, f^+car$^+$	Y^S and Y	0.64	55	0.059
(4)	X, y^+f^+car$^+$	Y^S and Y	0.55	80	0.371

All mating chambers were run with 12 males of each genotype and 24 X·Y^L/ /X·Y^L, $y f car$ females. Unless otherwise indicated males were also $y f car$. Where males were not phenotypically distinguishable, the tips of the wings were clipped to allow identification; different genotypes were clipped in replicate chambers. No significant effect of clipping was observed. n, Total number of matings observed. Goodness-of-fit χ^2 probabilities were corrected for continuity when the associated probability was greater than 0.10.
*$P<0.01$.
†$P<0.005$.
‡$P<0.05$.

acquired in the allopatric state[1-4]. One of the major problems of evolutionary genetics is, then, characterisation of the genetic differences which actually produce the reproductive and sexual isolation characteristic of species. A partial answer to this problem has been provided by estimating the degree of structural gene divergence over various taxonomic levels—populations, subspecies, semispecies, sibling species and morphologically distinguishable species[5,6]. The general picture emerging from these studies is moderate structural gene differentiation at the subspecies level. Once reproductive isolation has been achieved, little additional change seems to be required to establish sexual isolation between subspecies. There are, however, exceptions to this generalisation[7,9], and they do not seem to be infrequent. Speciation may occur as a result of chromosomal rearrangements, accompanied by little, if any, detectable change in the structural genes usually studied.

Attempting to explain the preponderance of sterility and inviability in the heterogametic progeny of hybrid matings, Haldane[11] suggested that sex chromosome translocations might be the cause. Tracey[12] estimated fixation probabilities

of an X·Y^L and a Y^S chromosome in laboratory populations of *Drosophila melanogaster*. The estimated probability of fixation exceeded neutral expectation for both of these translocation chromosomes; moreover X·Y^L//Y^S males had a mating advantage compared with X·Y^L//Y males when the females competed for were X·Y^L//X·Y^L. We have examined the genetic basis of this mating advantage and found that it is primarily determined by the Y^S chromosome, with the Y^L segment attached to the X and the *yellow* allele playing significant but lesser roles. In addition we have found that a population of *D. melanogaster* bearing X·Y^L, Y^S and the *yellow* allele exhibits significant sexual isolation with respect to a karyotypically normal strain homozygous for the eye mutant *sparkling-poliert*.

The X·Y^L chromosomes of all females used to test mating advantage carry the markers *yellow, forked* and *carnation*. Three Y chromosomes have been used: a normal Y, a Y^S which lacks the Y^L arm but is fertile when X·Y^L//Y^S and a Y^S bearing a translocated normal allele of *yellow* designated Y^S, y^+. Observation chambers[13] were used to study mating; the genotypes used were produced by crossing to

Table 2 Multifactor genetic basis of X·Y^L/ /Y^S male mating advantage

	X chromosomes (A)	(B)	Y chromosomes (A)	(B)	A mating frequency	n	χ^2 probability
(1)	X·Y^L	compared with X	Y^S	compared with Y^S, y^+	0.71	63	0.001*
(2)	X·Y^L	compared with X·Y^L	Y^S, y^+	compared with Y	0.71	147	0.001*
(3)	X·Y^L, y^+	compared with X·Y^L	Y^S	compared with Y	0.67	60	0.014†
(4)	X	compared with X	Y^S, y^+	compared with Y	0.76	80	0.001*
(5)	X	compared with X·Y^L	Y^S	compared with Y^S, y^+	0.31	48	0.014†
(6)	X	compared with X·Y^L	Y	compared with Y^S, y^+	0.09	53	0.001*
(7)	X, y^+f^+car$^+$ compared with X, f^+car$^+$		Y	compared with Y^S	0.93	30	0.001*

Male genotypes were combinations of columns A/ /A and B/ /B; as in Table 1, female genotypes were always X·Y^L/ /X·Y^L, $y f car$, and 24 females were used in each chamber. Twelve males of each genotype were used and they were $y f car$ except where wild alleles are noted. n, Total number of matings observed.
Goodness-of-fit χ^2 probabilities were corrected for continuity where the associated probability was less than 0.10.
*$P<0.005$.
†$P<0.05$.

a locally collected wild-type stock and selecting the desired recombinants.

Yellow males are known to mate less frequently than their normal counterparts[14,15]; therefore, we first compared *yellow* and normal bodied males (Table 1). There is no demonstrable difference in mating success between these males when they are fertile except for flies with normal chromosomes. X//Y males of normal body colour show a significant mating advantage. Similarly, the presence or absence of the Y^L segment on the X chromosome does not significantly affect mating success when this is the major genetic difference between competing males. Note, however, that the equivalence of *yellow* and normal male mating frequency does not hold when the males lack the Y^L segment and the Y^8 carries a y^+ marker. When Y^8- and Y-bearing males are compared, the Y^8 males do significantly better when they are fertile regardless of body colour. The absence of the Y^L segment on the X not only makes X//Y^8 males sterile, but also eliminates the mating advantage for flies of normal colouration. When *yellow* X//Y^8 and X//Y males competed, a nearly significant excess of X//Y^8 matings was observed. If this difference is real, it seems that the *yellow* allele and the Y^L segment interact with the Y^8 to produce the mating advantage observed in X·Y^L//Y^8 males.

Further evidence of interaction is presented in Table 2 which summarises mating chamber studies in which both the X and Y chromosomes of males differed. Comparison of studies 1 and 5 (Table 2), in which all Y chromosomes were Y^8 but the *yellow* marker and the X·Y^L chromosome were juggled, suggests that *yellow* males are more readily accepted by X·Y^L//X·Y^L females when the males carry the X·Y^L. In the presence of a normal X chromosome, the *yellow* mating advantage disappears. The comparisons of X·Y^L and X males (Table 1) indicate that this difference is attributable to a *yellow* by X·Y^L interaction. Males of normal body colour mate more frequently when they are of normal karyotype or carry a Y^8 marked with y^+ (Table 1, A5 and 6; Table 2, studies 2, 4, 5 and 6). The body colour equivalency apparent in Table 1, however, indicates that this advantage is an effect of the Y^8 rather than the y^+ allele. Moreover, X//Y^8 males are sterile; any mating advantage of such males is meaningless in an evolutionary context.

Thus, mating advantage of X·Y^L//Y^8 males is primarily affected by the Y^8 chromosome and augmented by a *yellow* –X·Y^L interaction. As X·Y^L//Y^8 is fitter than X//Y when females are X·Y^L//X·Y^L, it is important to assess the degree to which an X·Y^L–Y^8 system is sexually isolated from normal, X–Y populations of *D. melanogaster*. Conclusive tests of the degree of sexual isolation attributable to the translocation and to the *yellow* marker will require extensive homogenisation of genetic background. We have, however, obtained exciting results in a preliminary test. Males and females from the X·Y^L, *y f car*; Y^8 stock were placed in observation chambers with a wild strain containing the marker gene *sparkling–poliert* eye. The degree of homogametic mating exceeds 82%; 70 of 85 matings were between males and females of the same stock. This difference is highly significant and yields a sexual isolation coefficient equal to 0.65 ± 0.08; a value in excess of those obtained for populations and subspecies of *D. willistoni*[16], but within range of values obtained for semispecies of *D. paulistorum*[17].

To the extent that we may extrapolate from laboratory studies, these and previous[13] results suggest that speciation may occur with minimal genic change. The fixation of X·Y^L and Y^8 in an isolated population is a virtual certainty given sufficient time. In the absence of other change this X·Y^L–Y^8 population is reproductively isolated from its progenitor. The additional substitution of the *yellow* allele, which is itself advantageous in some conditions[14], engenders both a greater mating advantage and a rather high degree

of sexual isolation. In the light of the high frequency of sex chromosome polymorphisms, within and among species[19,20], and the low degree of genetic divergence between some of these species[7,9], Haldane's explanation of reproductive isolation requires serious consideration.

M. L. TRACEY
S. A. ESPINET

Department of Biological Sciences,
Brock University,
St Catharines, Ontario,
Canada L2S 3A1

1 Mayr, E., *Animal Species and Evolution* (Belknap Press/Harvard, Cambridge, Massachusetts, 1963).
2 Dobzhansky, T., *Genetics of the Evolutionary Process* (Columbia University Press, New York, 1970).
3 Lewontin, R. C., *The Genetic Basis of Evolutionary Change* (Columbia University Press, New York, 1974).
4 Nei, M., *Molecular Population Genetics and Evolution* (American Elsevier, New York, 1975).
5 Ayala, F. J., in *Evolutionary Biology*, 8 (edit. by Dobzhansky, T., Hecht, M. K., and Steere, W. C.), 1–78 (Plenum, New York, 1975).
6 *Molecular Evolution* (edit. by Ayala. F. J.), (Sinauer Associates, Sunderland, Massachusetts, 1976).
7 Gottlieb, L. D., *Evolution*, 27, 205–214 (1973).
8 Gottlieb, L. D., *Am. J. Bot.*, 60, 545–553 (1973).
9 Nevo, E., and Shaw, C. R., *Biochem. Genet.*, 7, 235–241 (1972).
10 Nevo, E., Kim, Y. J., Shaw, C. R., and Thaeler, C. S., *Evolution*, 28, 1–23 (1974).
11 Haldane, J B. S., *The Causes of Evolution* (Longmans, London, 1932).
12 Tracey, M. L., *Genetics*, 72, 317–333 (1972).
13 Elens, A A., and Wattiaux, J. M., *Drosophila Inf. Service*, 39, 118–119 (1964).
14 Bastock, M., *Evolution*, 10, 421–439 (1956).
15 Threlkeld, S F. H, Procwat, R. A., Abbott, K S, and Young, A. D., *Nature*, 247, 232–233 (1974).
16 Ayala, F. J., and Tracey, M. L, *J. Hered*, 64, 120–124 (1973).
17 Dobzhansky, T., and Pavlovsky, O., *Nature*, 230, 289–292 (1971).
18 Kalmus, H, *Proc. R. Soc.*, 130, 185–201 (1941).
19 Lewis, K. R., and John, B., *Chromosome Marker* (Churchill, London, 1963).
20 Hewitt, G. M., *Heredity*, 35, 375–387 (1975).

Lipoprotein of Gram-negative bacteria is essential for growth and division

THE lipoprotein of Gram-negative bacteria has become a cell surface component of central interest following its characterisation by Braun and coworkers[1]. This small protein (molecular weight 7,500) is unique in containing as its N-terminal amino acid a glycerylcysteine to which three fatty acids are covalently bound[2]. It is also unusual in its subcellular distribution: one-third of the total lipoprotein is covalently linked to the cell wall peptidoglycan while the remaining two-thirds exists in an unlinked or "free" form in the outer membrane[4]. Free and murein-linked lipoprotein seem to be identical in structure[4,6], and pulse–chase experiments indicate that free lipoprotein is the precursor to murein-linked lipoprotein[4]. Braun's lipoprotein or an immunologically related protein has been found in a wide variety of Gram-negative organisms[7,8], and in *Escherichia coli* is present in about 3×10^5 copies per cell[1]. These observations suggest that lipoprotein performs some essential function(s); however, the nature of that function is unknown. It has been suggested that lipoprotein may have a role in anchoring the outer membrane to the cell wall[5], or function in transporting small molecules through the outer membrane[9] or be involved in cell division[10,11]. To clarify the role of Braun's lipoprotein, we have isolated a mutant deficient in its synthesis.

The procedure used for isolation of the mutant was based on some of the unusual properties of the lipoprotein. The primary amino acid sequence of Braun's lipoprotein has been determined[12]; it has been found to lack five amino acids, including proline, histidine and tryptophan. Hirashima and Inouye[13] have demonstrated that during starvation for these amino acids, *E. coli* continues to synthesise lipoprotein, although synthesis of other proteins is drastically reduced. This observation suggested the possibility of enriching for a lipoprotein mutant by starving cells for proline, histidine and tryptophan, labelling with arginine, and performing a suicide selection. Since 42% of the total radioactivity incorporated by our strain

after amino acid starvation was found in lipoprotein (as compared to 3% in normal conditions), on prolonged storage a mutant which had not synthesised lipoprotein should survive longer than wild type. Two additional considerations taken into account were that lipoprotein might be essential for growth, and that the suicide would only enrich for mutants which either did not synthesise the protein or degraded it very rapidly. This prompted the use of a strain carrying a temperature-sensitive amber suppressor, which would allow the isolation of a conditional amber mutant in lipoprotein.

To perform the suicide selection, a strain which carried a temperature-sensitive suppressor and which could be starved for the appropriate amino acids was constructed by transducing D117 (an *E. coli* K12 strain of genotype *pro⁻his⁻trp⁻Su⁻Δlac/-F'lac⁻* amber) with bacteriophage P1 grown on *E. coli* MB188 (*trp⁻*am *SuIII*-A81 ts from K. Begg). A transductant was selected which was phenotypically *lac⁺trp⁺* at 30 °C and *lac⁻trp⁻* at 42 °C; the transductant also showed temperature sensitivity in its ability to allow growth of amber phage. It was cured of the episome by acridine orange. The resulting strain

Fig. 2 Effect of temperature shift on growth and viability of ST715. ST715 was grown at 30 °C in L broth[26] and at time 0 diluted 1/10 into prewarmed (42 °C) L broth. Viable counts (O) and absorbance (●) were measured. The solid line is a growth curve (absorbance) of the parent (M7) at 42 °C.
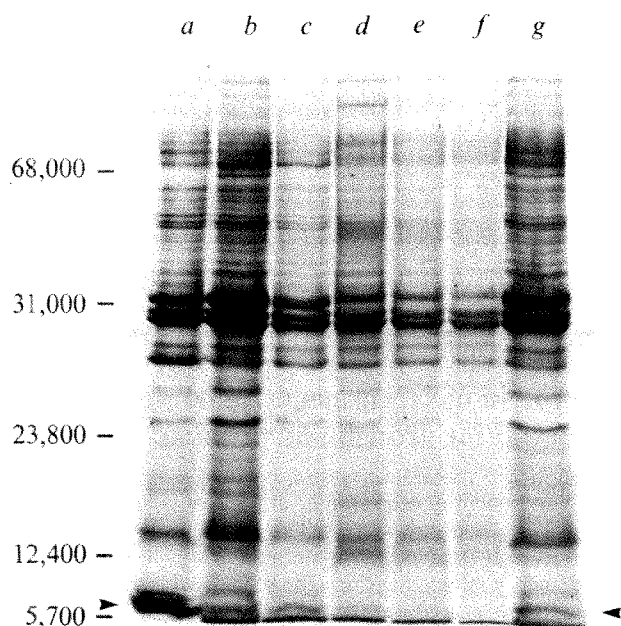
Fig. 1 SDS–acrylamide gel electrophoresis of envelope fractions from M7 and ST715 grown at 30 and 42 °C. M7 and ST715 were grown at 30 °C in minimal medium, filtered, and resuspended in medium containing cold arginine (2 µg ml⁻¹). A portion of each culture was shifted to 42 °C, and aliquots withdrawn after 30, 60, and 90 min. Each aliquot was labelled at 42 °C for 30 min by the addition of ¹⁴C-arginine (1 µCi ml⁻¹, specific activity 277 mCi mmol⁻¹). Control cultures (M7 and ST715 at 30 °C, M7 at 42 °C) were labelled 30 min after shift. Cultures were collected and envelope fractions prepared. These were dialysed overnight against 0.01 M Tris-HCl pH 6.8 at 4 °C, and collected by centrifugation at 90,000*g* for 60 min. Envelope fractions were solubilised in sample buffer[23] containing 0.5% SDS and applied to a 12.5% acrylamide slab gel[24] containing 0.5% SDS along with molecular weight markers. The gel was stained with Coomassi brilliant blue[25], dried and autoradiographed. Sample wells contain: *a*, M7 labelled at 30 °C, 60,000 c.p.m. applied; *b*, M7 labelled at 42 °C, 115,000 c.p.m. applied; *c*, ST715 at 30 °C, 40,000 c.p.m. applied; *d*, ST715 after 30 min. at 42 °C, 50,000 c.p.m. applied; *e*, ST715 after 60 min at 42 °C, 30,000 c.p.m. applied; *f*, ST715 after 90 min at 42 °C, 27,000 c.p.m. applied; *g*, same as *b*. Molecular weight standards were bovine serum albumin (68,000), DNase (31,000), trypsin (23,800), cytochrome *c* (12,400) and insulin (5,700). Free lipoprotein migrates immediately behind the front and is indicated by an arrow. The band of slightly higher molecular weight migrating behind free lipoprotein is presumably murein-linked lipoprotein which has been released by endogenous murein hydrolases. This assumption is based on the fact that neither band contains histidine (data not shown), and that the synthesis of both seems to decrease in a coordinate fashion during incubation of ST715 at 42 °C.
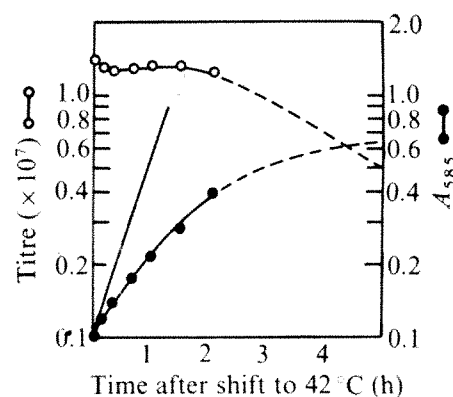
M7 (*pro⁻his⁻trp⁻*am *SuIII*-A81 ts), was mutagenised to 5% survival with nitrosoguanidine[14] diluted 1/10, and allowed to grow overnight at 30 °C in E medium[15] supplemented with the required amino acids. The culture was filtered, washed, resuspended in 50 ml E medium lacking proline, histidine, and tryptophan, and incubated at 42 °C for 60 min. ³H-arginine (50 µCi ml⁻¹, specific activity 28.7 Ci nmol⁻¹) was added, and incubation at 42 °C continued for 90 min. The cells were collected by filtration, washed with minimal salts containing unlabelled arginine (500 µg ml⁻¹), and then washed with cold water to remove pools of unincorporated radioactive arginine. The cells were resuspended in cold E salts and stored at 4 °C. Viability was measured at intervals by plating at 30 °C. The cell titre decreased exponentially, and had fallen from an initial value of 1 × 10⁸ ml⁻¹ to 3 × 10² ml⁻¹ after 45 d. Unlabelled control cells decreased in viability by less than 50% during this time. After 45 d, all surviving cells were plated on minimal plates and screened for temperature-sensitive growth on nutrient agar plates. Of 23,000 total survivors obtained from several experiments, 513 were temperature sensitive, and these were assayed for lipoprotein synthesis at 30 and 42 °C.

One of the temperature-sensitive survivors was found to synthesise reduced amounts of murein-linked lipoprotein at 42 °C (Table 1). Murein-linked lipoprotein accounts for 3.2% of the arginine incorporated into the envelope fraction of M7 at 30 °C; this figure is slightly higher at 42 °C. In contrast, strain ST715, although showing normal lipoprotein levels at

Fig. 3 Effect of temperature shift on morphology of ST715. ST715 was grown in L broth at 30 °C (*a*), and shifted to 42 °C for 3 h (*b*). Cells were collected by centrifugation, resuspended on 0.01 M Tris pH 6.8, and stained with 2% uranyl acetate. Micrographs were taken in a Joelco 100B electron miscroscope at a magnification of × 5,000.

Table 1 Assay of murein-linked lipoprotein in parent and mutants grown at 30 and 42 °C

		Total envelope (^{14}C-Arg c.p.m.)	Lipoprotein (c.p.m.)	Lipoprotein/total envelope ratio (%)
M7	30 °C	73,968	2,392	3.2
	42 °C	77,536	3,304	4.3
ST715	30 °C	39,276	1,255	3.2
	42 °C	80,055	636	0.8

This assay is based on the observation[20] that murein-linked lipoprotein can migrate on chromatograms in an isobutyric acid–ammonia solvent after it is released from murein by lysozyme digestion. ST715 and M7 were grown at 30 °C in E medium supplemented with required amino acids and arginine. In the early phase of exponential growth, the culture was filtered and cells resuspended at 30 °C in medium containing arginine (4 μg ml^{-1}). After 10 min, a portion of each culture was shifted to 42 °C and incubated for 30 min; control cultures remained at 30 °C. ^{14}C-arginine (specific activity 277 mCi mmol^{-1}) was added to 0.5 μCi ml^{-1}, and incubation continued for 30 min. Cultures were chilled, collected by centrifugation at 4 °C, washed once with cold saline and re-suspended in 1 ml cold 0.01 M Tris HCl *p*H 6.8. The cells were ruptured by sonication, centrifuged for 5 min at 3,000*g* to remove unbroken cells, and then centrifuged at 90,000*g* for 60 min to collect envelopes. These were resuspended in a small volume of water and boiled for 5 min. A small aliquot was spotted on Whatman 3MM chromatography paper for direct counting; the remaining sample was applied in two equal portions to Whatman 3MM paper and chromatographed overnight in isobutyric acid–1 N NH$_4$OH (5:3) (descending chromatography). This procedure separates murein-linked lipoprotein (origin) from free lipoprotein and other envelope proteins which migrate with an R_f of 0.8. Chromatograms were dried, and a 2-cm region around the origin cut out. One of each duplicate sample was incubated overnight at 37 °C in 2 ml 0.5 N ammonium acetate containing egg white lysozyme 500 μg ml^{-1}; the control was incubated in buffer alone. Incubation with lysozyme releases the lipoprotein from the cell wall material at the origin but does not liberate the lipoprotein from the paper.

The origins were washed twice with water, dried, sewn on to Whatman 3MM paper, and rechromatographed in isobutyric acid–1N NH$_4$OH (5:3) for 4 h. Murein-linked lipoprotein migrates with an R_f of 0.8 during this chromatography. Chromatograms were dried, cut into 2-cm pieces and counted. Each value was corrected for counts released during the control incubation (from 10–20% of the value obtained using lysozyme), and samples were normalised by comparing each lipoprotein value to total envelope counts applied. This assay is specific for murein-linked lipoprotein since there is little release of radioactivity in the absence of lysozyme; the released material contains palmitic acid and diaminopimelic acid but not histidine; when eluted from the chromatogram and subjected to SDS–acrylamide gel electrophoresis, all the material comigrates with *in vivo*-labelled murein-linked lipoprotein (data not shown).

30 °C, shows a fourfold decrease in incorporation into lipoprotein after 60 min at 42 °C. A second survivor also showed a twofold decrease in murein-linked lipoprotein at 42 °C as well as slightly reduced levels at 30 °C (data not shown); this strain was very unstable and further studies were performed on ST715 only.

To determine whether ST715 also showed a temperature-dependent reduction in synthesis of free lipoprotein, M7 and

ST715 were incubated at 30 and 42 °C for various periods of time and then labelled with ^{14}C-arginine. Whole envelope fractions were prepared and subjected to sodium dodecyl sulphate (SDS)–acrylamide gel electrophoresis as shown in Fig. 1. This gel does demonstrate a decrease in synthesis of lipoprotein in ST715 as a function of time of incubation at 42 °C; this decrease is quite gradual, and some synthesis is still detectable after 2 h at 42 °C, presumably reflecting leakiness of the temperature-sensitive suppressor. Other changes in the envelope fraction of ST715 at 42 °C can be seen; notably the appearance and gradual disappearance of a protein band of high molecular weight ($\sim$ 90,000) and the disappearance of a band of $\sim$ 70,000. The relationship of these changes to the lipoprotein lesion is unclear and will be studied in revertants.

The effects of lipoprotein deficiency on cell physiology were studied by shifting an exponentially growing broth culture of ST715 from 30 to 42 °C (Fig. 2). This shift results in an abrupt cessation in the increase of viable counts. Cell mass continues to increase, however, and filaments are formed (Fig. 3). After several hours at 42 °C, viability begins to decrease slowly; at this time cells are 4–8 times their normal length. Thus, continued lipoprotein synthesis seems necessary for cell division.

To determine whether lipoprotein deficiency, filament formation and temperature-sensitive growth were the result of a single mutation, revertants which could grow at the non-permissive temperature were selected. These were found to occur at a frequency of approximately 10^{-9}. Ten revertants were selected for further study, and all were found to have increased levels of murein-linked lipoprotein at 42 °C. Two revertants, E4 and E10, which showed normal growth rates at 42 °C, were analysed for murein-linked lipoprotein and found to have levels comparable to M7 (Table 2). Interestingly, E10 seems to be a wild-type revertant, whereas E4 has apparently reverted to temperature-independent suppression. The fact that activation of a suppressor at 42 °C enables E4 to grow and make lipoprotein at this temperature suggests that ST715 carries an amber mutation. Other revertants studied seemed to be partial revertants, which grew with a greater than normal generation time at 42 °C, and, showed only partial restoration of bound lipoprotein levels at 42 °C (ranging from a two- to fourfold increase over mutant levels in an experiment in which the parental level was 7 times that of ST715). This may indicate that lower amounts of lipoprotein than observed in the parent are sufficient to support growth. The results shown in Table 2 also demonstrate that growth at 42 °C, filament formation, and lipoprotein content are properties linked by transduction. These results indicate that all properties of the mutant are due to a single mutation, and suggest that lipoprotein may be necessary for normal growth and division.

Wu and Lin[21], using a suicide technique similar to that described here, have recently isolated a mutant deficient in linking lipoprotein to murein. Their mutant, like ST715, shows temperature sensitivity in growth and division. Since the strain

Table 2 Properties of revertants and transductants of ST715

Strain	Growth at 42 °C		Plates T4amB25		Morphology		Murein-linked lipoprotein (at 42 °C)/ total envelope ratio
	+ tryp.	− tryp.	30 °C	42 °C	30 °C	42 °C	
M7	+	−	+	−	rod	rod	4.5%
ST715	−	−	+	nd	rod	fil	0.5%
E4	+	+	+	+	rod	rod	4.0%
E10	+	−	+	−	rod	rod	3.1%
TD4	+	−	+	−	rod	rod	5.3%
TD6	+	−	+	−	rod	rod	3.6%

Revertants of ST715 were isolated by spreading approximately 3×10^9 cells on minimal plates and incubating at 39 °C. Revertant clones were purified, tested for amino acid requirements and streptomycin resistance, and grown in minimal medium at 42 °C to determine growth rate. Murein-linked lipoprotein was measured on cultures which had been pregrown at 30 °C, shifted to 42 °C for 60 min, and then labelled at 42 °C for 30 min with ^{14}C-arginine. Growth requirement for tryptophan was determined on plates; phage sensitivity was measured by plating efficiency at 30 and 42 °C; and morphology examined by phase microscopy.
P1 transduction was performed as described[27] using *E. coli* K12 strain GR2131 (genotype *pro-leu-trp-gal-man-xyl-mal-lac-phoA purE* from G. Willsky) as a donor. Transductants which were able to grow at 42 °C were isolated, purified, and tested for amino acid requirements. Two transductants, TD4 and TD6, were assayed for murein-linked lipoprotein and tested for other properties as described above.

isolated by Wu and Lin seems, however, markedly deficient in linking lipoprotein to murein at the permissive temperature, it is possible that murein-linked lipoprotein does not have an essential role in growth and division of *E. coli*. This would imply that it is the deficiency in free lipoprotein which is responsible for the phenotype of ST715. In addition, until it can be shown that ST715 carries a lesion in the structural gene for lipoprotein, it remains possible that ST715 is affected in growth and lipoprotein through a mutation which affects both properties indirectly. An example would be a protease responsible for processing lipoprotein and other essential proteins.

Mutants with reduced levels of other major outer membrane proteins (of molecular weight 30,000–40,000) have been described[14,15] as well as mutants which lack one[16,17] or all[18] of these proteins; however, such mutants are viable and morphologically normal. This is in contrast to ST715, which is unable to grow and divide at the non-permissive temperature when lipoprotein formation is greatly reduced. Braun's lipoprotein thus may serve a more vital function(s) than do these other major outer membrane proteins. Further studies with the lipoprotein mutant described here should define the role of this major component of the outer membrane in cell growth and division.

This work was supported by a grant from the NIH. We thank E. Ekmejian, O. Fields and D. Dealy for assistance.

Suzy V. Torti
James T. Park

Department of Molecular Biology and Microbiology,
Tufts University Medical School,
136 Harrison Avenue,
Boston, Massachusetts 02111

1 Braun, V., *Biochim. biophys. Acta*, **415**, 335–377 (1975).
2 Hantke, K., and Braun, V., *Eur. J. Biochem.*, **34**, 284–296 (1973).
3 Braun, V., and Rehn, K., *Eur. J. Biochem.*, **10**, 426–438 (1969).
4 Inouye, M., Shaw, J., and Shen, C., *J. biol. Chem.*, **247**, 8154–8159 (1972).
5 Hirashima, A., Wu, H. C., Venkateswaran, P. S., and Inouye, M., *J. biol. Chem.*, **248**, 5654–5659 (1973).
6 Braun, V., Hantke, K., and Henning, U., *FEBS Lett.*, **60**, 26–28 (1975).
7 Braun, V., Rehn, K., and Wolff, H., *Biochemistry*, **9**, 5041–5049 (1970).
8 Halegoua, S., Hirashima, A., and Inouye, M., *J. Bact.*, **120**, 1204–1208 (1974).
9 Inouye, M., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 2396–2400 (1974).
10 Burdett, I. D. J., and Murray, R. G. E., *J. Bact.*, **119**, 1039–1056 (1974).
11 James, R., and Gudas, L. J., *J. Bact.*, **125**, 374–375 (1976).
12 Braun, V., and Bosch, V., *Proc. natn. Acad. Sci. U.S.A.*, **69**, 970–974 (1972).
13 Hirashima, A., and Inouye, M., *Nature*, **242**, 405–407 (1973).
14 Adelberg, E., Mandel, M., and Chen, G., *Biochem. biophys. Res. Commun.*, **18**, 788–795 (1965).
15 Vogel, H., and Bonner, D., *J. biol. Chem.*, **218**, 97–106 (1956).
16 Ames, G. F., Spudich, E. N., and Nikaido, H., *J. Bact.*, **117**, 406–416 (1974).
17 Koplow, J., and Goldfine, H., *J. Bact.*, **117**, 527–543 (1974).
18 Chai, T., and Foulds, J., *J. molec. Biol.*, **85**, 465–474 (1974).
19 Skurry, R. A., Hancock, R., and Reeves, P., *J. Bact.*, **119**, 726–735 (1974).
20 Henning, U., and Haller, I., *FEBS Lett.*, **55**, 161–164 (1975).
21 Wu, H., and Lin, J., *J. Bact.*, **126**, 147–156 (1976).
22 Braun, V., and Bosch, V., *FEBS Lett.*, **34**, 302–306 (1973).
23 Laemmli, U. K., *Nature*, **227**, 680–685 (1970).
24 Studier, F. W., *J. molec. Biol.*, **79**, 237–248 (1973).
25 Willsky, G. R., and Malamy, M. H., *J. Bact.*, **127**, 595–609 (1976).
26 Lennox, E. S., *Virology*, **1**, 190–206 (1955).
27 Willsky, G. R., Bennett, R. L., and Malamy, M. H., *J. Bact.*, **113**, 529–539 (1973).

Prevention of the contagious spread of feline leukaemia virus and the development of leukaemia in pet cats

Clustering of cases of feline lymphosarcoma (LSA) or leukaemia has been observed by veterinarians for many years[1-4]. The first anti-feline leukaemia virus (FeLV) serum was reported in 1969[5] and by 1970 a simple indirect immunofluorescent antibody (IFA) test for the detection of FeLV in the peripheral blood of infected cats had been developed[6]. Using this test we found that 33% of the healthy pet cats exposed to cats with FeLV-associated diseases were infected with FeLV. All the cats known not to have been exposed to cats with FeLV-associated diseases were uninfected and only 0.31% of the stray cats, with an unknown history of FeLV exposure, were infected[3]. Once persistently viraemic with FeLV, most

cats remain infected for their entire lives. These results showed conclusively that FeLV is transmitted between cats by infection or contagion (that is, horizontally), in contrast to the oncornaviruses of inbred mice which are predominantly transmitted genetically from the parents to the offspring by means of the chromosomes (that is, vertically)[7].

FeLV causes four fatal diseases in pet cats. The diseases are: (1) LSA[4,5], (2) non-regenerative anaemia[8,9], (3) a panleukopaenia-like syndrome[3], and (4) thymic atrophy[10]. FeLV is also associated with, but not yet proven to be the cause of, various myeloproliferative disorders[11] and foetal abortions and resorptions (F. Goldsmith and W. D. H., Jr, unpublished)[12,13]. Healthy FeLV-infected cats have a greatly increased chance of developing one of the FeLV-related diseases or of developing secondary diseases due to the immunosuppressive effects of FeLV[12,14].

The spread of an infectious disease agent may be prevented by various methods. These are: (1) vaccination, (2) killing the infectious agent in the environment by disinfection, and (3) removal of either the infected sick animal, the healthy carrier host or the vector(s). In veterinary medicine, when vaccination is not available, removal of the carrier host is commonly used to prevent the spread of infectious disease agents[15].

Our earlier epidemiological studies showed that FeLV is an infectious agent for pet cats[3]. The ultimate goal of any epidemiological study is to prevent the spread of a disease by identifying the aetiological agent so that it can be removed from the environment. We report here the successful use of an FeLV test and removal programme to prevent the spread of FeLV, and the diseases it causes, among pet cats. The FeLV test and removal programme consists of the identification and removal, by euthanasia or isolation, of FeLV-infected healthy cats from contact with uninfected cats[16]. The IFA test for FeLV detects FeLV gs antigens in the peripheral blood leukocytes and platelets of infected cats[3,6] and we have previously reported that a positive IFA test result means that the cat is viraemic[6,13].

In most multiple cat households in this study an index FeLV test was performed on a sick or healthy cat to establish the FeLV status of the households. The index test-positive cat was removed and all the remaining healthy cats in the household were then tested for FeLV (initial test). In the programme, any exposed healthy cats which were found to be infected with FeLV were removed immediately. After the initial test of all the healthy cats, the household was quarantined so that no cats were allowed to leave and no new cats were brought into the household. After a 3-month interval a second test for FeLV was performed on all the remaining healthy cats in the household and any secondarily infected cats were removed. Cats which were FeLV uninfected in the initial test but which were found to be FeLV infected in the second test will be referred to as secondarily infected cats in this paper. Two IFA tests for FeLV are required for exposed cats because the natural incubation period for the establishment of viraemia can be as long as 3 months, and there is a possibility that an FeLV infection, occurring just before the first test, would not be detected in that test[13]. If any healthy cats were positive in the second IFA test they were removed from the household and a third test was conducted 3 months later. When all the cats remaining in a household after the infected cats had been removed were negative in two consecutive IFA tests, conducted 3 months apart, the household was considered FeLV free. All new cats which were introduced into an FeLV-free household were tested for FeLV before they were allowed to have contact with the uninfected cats.

A total of 1,260 healthy cats, from 65 households in which cats had developed FeLV-related diseases and from 11 breeding catteries, have been studied (Table 1). In 51 of these households the test and removal programme was implemented, while in 25 households the owners chose not to implement the programme. These 25 households served as controls and were therefore compared with the 51 households in which the

Table 1 Disease and FeLV status of household groups before the FeLV test and removal programme

Household groups	Number of households	Number of cats with FeLV diseases				Index FeLV test*		Total cats remaining after diseased cats died
		Total cats	Lymphosarcoma	FeLV anaemia	Other	Household status FeLV infected	Not tested	
Households where FeLV-infected healthy cats were removed	51	923	42	6	28	37	14†	847
Households where FeLV-infected healthy cats were not removed	25	472	37	4	18	23	2‡	413
								Total 1,260

*An index FeLV test was the first IFA test of a cat in a household.
†Index FeLV tests were not performed in 10 breeding catteries because no cats had previously developed an FeLV disease. No index tests were performed in the other 4 households where cats had developed FeLV diseases because these cats had died before our study began.
‡An index FeLV test was not performed in 1 household (a breeding cattery) because no cat had previously developed an FeLV disease. An index test was not performed in the other household where a cat had developed an FeLV disease because this cat had died before our study began.

Table 2 Summary of the test and removal programme to prevent the contagious spread of FeLV in healthy pet cats

Household groups	Number of households	Number of healthy cats in households	Number of healthy cats tested	Initial test of healthy cats FeLV results		Initial infection rate	Number of healthy cats removed	Second test* of healthy cats FeLV results		Number of healthy cats which became secondarily† infected	Secondary infection rate of healthy cats
				Infected	Uninfected			Infected	Uninfected		
Households where FeLV-infected healthy cats were removed	51	847	817	190	657	22.4%	190	3	654	3	0.46%
Households where FeLV-infected healthy cats were not removed	25	413	413	129	284	31.2%	0	184	229	55	19.3%

*The second FeLV test was performed at least 3 months after the initial FeLV test.
†A secondarily infected cat is one which was FeLV uninfected in the initial test but was found to be infected in the second test.

Table 3 Summary of disease development in household groups

Household groups	Number of households	Number of initial test FeLV-uninfected cats	Number of healthy cats which became secondarily infected	Disease development in secondarily* FeLV-infected healthy cats						Natural mortality rate in first test FeLV-uninfected§ healthy cats	Number of cats which died naturally	Natural mortality rate in secondarily FeLV-infected healthy cats
				Lymphosarcoma	Other FeLV† caused diseases	Other non-FeLV‡ caused diseases	Number of cats which died naturally	Alive	Euthanised			
Households where FeLV-infected healthy cats were removed	51	657	3	—	—	—	—	3	—	—	—	—
Households where FeLV-infected healthy cats were not removed	25	284	55	7	6	11	24	22	9	8.5%	24	52%

*A secondarily infected cat is one which was FeLV uninfected in the initial test but was found to be infected in the second test.
†4 cats with FeLV non-regenerative anaemia; 2 cats with the panleukopaenia-like syndrome.
‡3 cats with feline infectious peritonitis; 8 other diseases—these diseases occur frequently in FeLV-infected cats due to the immunosuppressive effects of the virus.
§These cats subsequently became secondarily infected with FeLV.

owners implemented the FeLV test and removal programme. A total of 847 cats were tested for FeLV in the 51 households in which the programme was implemented (Table 2). Of these 847 cats, 190 cats were found to be infected in the initial FeLV test (an initial infection rate of 22.4%) and were immediately removed from contact with the remaining 657 uninfected cats. Only 3 of these 657 uninfected healthy cats were found to be infected in the second test. In the 3-month period between the two IFA tests, 99.54% of the initially uninfected cats therefore remained uninfected (a secondary infection rate of 0.46%).

In the 25 households where the FeLV-infected healthy cats were not removed, and thus remained in continual contact with other cats in the households, a total of 413 cats were tested for FeLV and 129 were found to be infected in the initial test (an initial infection rate of 31.2%). Of the remaining 284 uninfected healthy cats in these households, 55 were found to be infected in the second test (a secondary infection rate of 19.3%) (Table 2). Thus only 80.7% of the initially uninfected cats remained uninfected in the 3-month period between the two tests. The secondary FeLV infection rate in the households in which the FeLV test and removal programme was not implemented was 42 times greater than that in households in which the programme was implemented. This difference is highly significant ($P \leqslant 0.0001$) by the χ^2 test. Thus the natural contagious spread of FeLV between cats was effectively prevented by the FeLV test and removal programme.

We have reported previously that FeLV-infected cats have a significantly higher chance of developing one of the FeLV-related diseases[3] and a much higher mortality rate than uninfected cats[17]. There was a marked difference in the occurrence of the disease and the mortality rate of cats living in the two household groups. The development of FeLV-related diseases was prevented in healthy cats living in the 51 households in which the test and removal programme was implemented. In contrast, in the 25 households where the programme was not implemented, 7 cats have developed LSA, 4 cats have developed FeLV non-regenerative anaemia and 13 cats have developed other FeLV-associated diseases within the 2-yr observation period (Table 3).

The natural mortality rate for the 3 healthy initially uninfected cats which became secondarily infected in those households in which the programme was implemented was not determined since the cats were euthanised immediately after they were found to be infected and thus did not have a chance to develop disease. In the 25 households in which the programme was not implemented 24 of the 55 secondarily infected cats subsequently died resulting in a mortality rate of 52%. These 24 deaths represent a natural mortality rate of 8.5% of the 284 initially FeLV-uninfected healthy cats (Table 3).

Some healthy cats living in households where the control programme was not implemented remained uninfected even though they were continually exposed to viraemic cats. To ensure that the test and removal programme, rather than FeLV neutralising antibody, was responsible for the protection of uninfected cats, cats living in selected households which implemented the programme were examined for neutralising antibody. Forty-one per cent of these cats tested had protective ($\geqslant 1:10$) neutralising antibody titres. Fifty-nine per cent of the cats thus had non-protective antibody titres and were susceptible to FeLV infection. Because of the implementation of the test and removal programme, however, these susceptible cats were not continuously exposed to FeLV and thus remained uninfected and did not develop any FeLV related diseases.

Pet cats of all ages are susceptible to FeLV infection in their natural household environments. The median age when FeLV was first detected in 49 of the secondarily infected pet cats in our study was 24.0 months. The median age at which these cats were last tested and found to be uninfected was 18.0 months. Therefore, between 18 and 24 months these adult pet cats became infected with FeLV in their household environments.

The FeLV test and removal programme is currently being used by many veterinarians to protect FeLV-uninfected cats

from FeLV infection and disease development[16]. The diseases caused by, or associated with, FeLV can be prevented by the test and removal programme. This is significant since diseases caused by and associated with FeLV are probably the most common killer of pet cats.

The results obtained by identifying and removing the healthy inapparent FeLV-infected carrier cats show for the first time that it is possible to prevent the contagious spread of a mammalian oncornavirus in the natural environment. FeLV and the diseases it causes may ultimately be controlled by a combination of the test and removal programme and vaccination.

W. D. HARDY, JR
A. J. MCCLELLAND
E. E. ZUCKERMAN

Memorial Sloan–Kettering Cancer Center,
New York, New York 10021

P. W. HESS

The Henry Bergh Memorial Hospital of the American Society
for Prevention of Cruelty to Animals,
New York, New York 10028

M. ESSEX

Department of Microbiology,
Harvard School of Public Health,
Boston, Massachusetts 02115

S. M. COTTER

Angell Memorial Animal Hospital,
Boston, Massachusetts 02115

E. G. MACEWEN
A. A. HAYES

The Animal Medical Center,
New York, New York 10021

1 Schneider, R., Frye, F. L., Taylor, D. O. N., and Dorn, C. R., *Cancer Res.*, 27, 1316–1322 (1967).
2 Brodey, R. S., McDonough, S., Frye, F. L., and Hardy, W. D., Jr, in *Comparative Leukaemia Research* (edit. by R. M. Dutcher), 333–342 (Karger, Basel, 1970).
3 Hardy, W. D., Jr, Old, L. J., Hess, P. W., Essex, M., and Cotter, S. M., *Nature*, 244, 266–269 (1973).
4 Jarrett, W. F. H., Crawford, E. M., Martin, W. B., and Davie, F., *Nature*, 202, 567–569 (1964).
5 Hardy, W. D., Jr, et al., *Science*, 166, 1019–1021 (1969).
6 Hardy, W. D., Jr, Hirshaut, Y., and Hess, P., in *Unifying Concepts of Leukaemia* (edit. by R. M., Dutcher, and L. Chieco-Bianchi), 778–799 (Karger, Basel, 1973).
7 Huebner, R. J., and Todaro, G. J., *Proc. natn. Acad. Sci. U.S.A.*, 64, 1087–1094 (1969).
8 Hoover, E. A., Kociba, G. J., Hardy, W. D., Jr, and Yohn, D. S., *J. natn. Cancer Inst.*, 53, 1271–1276 (1974).
9 Mackey, L., Jarrett, W., Jarrett, O., and Laird, H. M., *J. natn. Cancer Inst.*, 54, 209–217 (1975).
10 Anderson, L. J., Jarrett, W. F., Jarrett, O., and Laird, H. M., *J. natn. Cancer Inst.*, 47, 807–817 (1971).
11 Herz, A., Theilen, G. H., Schalm, O. W., and Munn, R. J., *J. natn. Cancer Inst.*, 44, 339–348 (1970).
12 Cotter, S. M., Hardy, W. D., Jr, and Essex, M., *J. Am. vet. Med. Ass.*, 166, 449–454 (1975).
13 Hardy, W. D., Jr, et al., *Cancer Res.*, 36, 318–324 (1976).
14 Essex, M., Sliski, A., Cotter, S. M., Jakowski, R. M., and Hardy, W. D., Jr, *Science*, 190, 790–792 (1975).
15 Blood, D. C., and Henderson, J. A., *Veterinary Medicine* (Williams and Wilkins, Baltimore, 1963).
16 Hardy, W. D., Jr, McClelland, A. J., Hess, P. W., and MacEwen, E. G., *J. Am. Animal Hosp. Ass.*, 10, 367–372 (1974).
17 Essex, M., Hardy, W. D., Jr, Cotter, S. M., Jakowski, R. M., and Sliski, A., *Infect. Immun.*, 11, 470–475 (1975).

Strain differences in the autoimmune response of mice to acetylcholine receptors

EXPERIMENTAL autoimmune myasthenia gravis (EAMG) has been induced in rabbits[1-4] guinea pigs[5,6] rats[4,6] and monkeys[7] by injection of the acetylcholine receptor (AChR) isolated from electric fish. The experimental disease, together with the demonstration of humoral[8-10] and cellular[11] immune responses to AChR in patients with myasthenia gravis (MG), verifies hypotheses[12,13] about the immunopharmacological block of AChR in MG. An experimental model disease in mice is needed for studies of the genetic aspects of MG and the role of the thymus as a specific antigen target, and as a source for immunocompetent helper and suppressor cells. We describe here the induction of EAMG in several inbred strains of mice and demonstrate different susceptibility to the disease in strains representing different haplotypes of the major histocompatibility complex (H-2). We describe some humoral and cellular aspects of the disease and demonstrate that in a susceptible mouse strain both humoral and cellular immune response to self-AChR are elicited. Finally we show that AChR behaves as a thymus-dependent antigen.

Mice were injected twice with a 9-week interval with 10 μg of purified *Torpedo californica* AChR (specific activity 10 nmol of toxin per mg protein) per mouse. The immunogen was emulsified in complete Freund's adjuvant (CFA) and injections were made into the footpads. Ten days or more after the second injection clinical signs of EAMG were observed in mice of the following strains: A/J and B10.A with the H-2ᵃ haplotype, C57BL/6, C3H.SW and CWB with H-2ᵇ, BALB/c and DBA/2 with H-2ᵈ, AKR/Cu and CKB with H-2ᵏ. The disease was not found in any mice with H-2ᵠ and H-2ˢ haplotypes (Table 1). The sick mice suffered from weight loss and exhibited signs of fatigue, hypoactivity, ruffled fur, paralysis of the limbs and motor impairment, which was accentuated by exercise. Their heads were sinking and their backs were exaggeratedly humped. Severely sick animals died from the disease, whereas in some cases the disease seemed to be transient.

Table 1 EAMG incidence and antibody titres against AChR in different strains of mice

Strain	H-2 Haplotype	log₂ Haemagglutination titre			Animals with clinical signs of EAMG
		1 °C day 21	1 °C day 63	2 °C day 8	
A/J	a	8.2	13.5	20.3	1/5
B10.A	a	7.5	11.2	20.0	2/5
C57BL/6	b	8.0	11.8	18.9	8/12
C3H.SW	b	7.5	11.0	19.3	2/5
CWB	b	7.5	11.5	19.6	2/10
BALB/c	d	7.5	11.2	19.8	2/5
DBA/2	d	8.0	11.0	20.2	2/6
C3H/Hej	k	7.5	12.0	18.5	2/5
CKB	k	7.6	11.3	20.0	2/6
AKR/Cu	k	7.8	11.4	19.0	6/6
DBA/1	q	7.2	11.3	20.2	0/6
SWR	q	8.2	11.4	20.6	0/6
SJL/J	s	8.0	11.8	21.0	0/6
ASW	s	7.3	11.7	20.6	0/6

Antibody titres were determined by micropassive-haemagglutination technique. Formalinised sheep red blood cells were coated with *Torpedo* AChR (100 μg per ml packed cells) by means of tannic acid.

The symptoms of mice with EAMG were reversed temporarily soon after intravenous injection of 5 μg of edrophonium chloride (Tensilon). In particular there was improvement of motor performance such as gripping and walking. Further pharmacological evidence for a physiological aberration in neuromuscular transmission in mice with EAMG was obtained from a curare ((+)-tubocurarine) test[14]. All the AChR-injected mice of strains in which symptoms of EAMG were observed, including individual mice which did not exhibit symptoms, had a higher sensitivity to curare than did the resistant strains. For example, injection of 0.6 nmol of curare intravenously was lethal for the AChR-injected C57BL/6 mice, tested 5 months after the first injection, whereas 2.4 nmol was required to kill

Table 2 Incorporation of tritiated thymidine by cultures of lymph node cells from C57BL/6 mice sensitised with AChR

Antigen added to sensitised lymph node cells (μg per culture)		Stimulation index*	
		day 10	day 42
AChR	(0.50)	9.24	5.16
	(0.25)	10.07	7.14
Mouse (C57BL/6)	(5.0)	4.30	Not done
muscle extract	(2.5)	2.58	Not done
Rabbit muscle extract	(5.0)	6.50	8.66
Lysozyme	(2.0)	1.16	0.88

Microculture lymphocyte transformation technique was used according to the method described by Oppenheim *et al.*[15]. Mice were injected with 10 μg of AChR in CFA in the footpads 10 or 42 d before the experiments.
*Ratio of c.p.m. of experimental to control cultures; average of five experiments.

Fig. 1 Electromyographs from myasthenic and normal C57BL/6 mice. Supramaximal tetanic stimulation was applied to the sciatic nerve at a frequency of 20 per s with rectangular pulses of 0.1 ms duration and 0.75–1.5 V amplitude. The evoked compound muscle action potential was recorded with a coaxial needle electrode embedded in the calf muscle by an HP electromyograph. *a*, Myasthenic mouse; *b*, control mouse.

Table 3 Antibody response to AChR in an adoptive cell transfer experiment in C57BL/6 mice

Group no.	Cells transferred into lethally-irradiated mice	AChR injected	log₂ Haemagglutination titre* day 7†	day 12	day 18
(1)	Normal spleen cells (30 × 10⁶)	10 µg in CFA	2.25	5.0	6.37
(2)	Normal spleen cells treated with ATS‡ (30 × 10⁶)	10 µg in CFA	1.20	1.66	1.75
(3)	Primed spleen cells§ (30 × 10⁶)	10 µg in CFA	ND	13.0	14.75
(4)	Primed spleen cells treated with ATS (30 × 10⁶)	10 µg in CFA	ND	1.70	1.88

Cell-transfer experiments into lethally irradiated mice (800 rad ⁶⁰Co whole-body irradiation) were performed in a procedure similar to that described by McDevitt and Tyan[17]. ND, Not done.
*Average titre of 10–25 mice.
†Days after injection of AChR.
‡Anti-thymocytic serum.
§Cells obtained from mice injected with 10 µg of AChR in CFA 3 weeks earlier.

normal C57BL/6 mice. On the other hand, AChR-injected SWR and DBA/1 mice (H-2�q) as well as SJL/J and ASW mice (H-2ˢ) which showed no symptoms were not significantly more sensitive to curare than normal siblings which had not received antigen. Neurophysiological studies of C57BL/6 mice, made with an HP electromyograph revealed a decrease in amplitude during repetitive nerve stimulation in sick mice (Fig. 1a) and a normal response in control animals (Fig. 1b). Moreover, AChR-injected C57BL/6 mice which had no symptoms also showed a decremental response to repetitive nerve stimulation. The decremental electromyographic response as well as the pharmacological effects of Tensilon and curare are common criteria for verifying the neuromuscular block in MG.

The autoimmune response against self-AChR in C57BL/6 mice injected with *Torpedo* AChR was demonstrated by cellular and humoral reactivity towards syngeneic muscle extracts. Cellular sensitivity was tested by the *in vitro* lymphocyte transformation technique. As Table 2 shows, lymph node cells of mice injected once with *Torpedo* AChR were stimulated *in vitro* when incubated with the immunogen, as well as with xenogeneic (rabbit) and syngeneic (C57BL/6) muscle extracts. Humoral autoimmune response was demonstrated by the ability of mouse anti-AChR sera to bind to syngeneic AChR. An AChR-rich fraction was prepared from C57BL/6 mouse muscle, by solubilisation with Triton in a procedure similar to that described by Lindstrom *et al.*[16] for the preparation of rat muscle extract. AChR in the extract was labelled with ¹²⁵I-bungarotoxin and the binding of labelled receptor to mouse antiserum was measured by radioimmunoassay. Sera from C57BL/6 mice immunised with *Torpedo* AChR bound, per ml of serum, an amount of mouse AChR corresponding to 0.35 pmol of toxin-binding sites.

Humoral immune response towards *Torpedo* AChR in the injected mice was determined using the micropassive-haemagglutination technique (Table 1). Significant titres of antibodies against AChR were observed in the mouse sera 3 weeks after a single immunisation. The antibody levels increased during the following 6 weeks and an additional significant increase was observed 8 d after the booster. No correlation between the antibody titres and incidence of the disease was found, as mice of all tested strains gave similar antibody titres (Table 1). It is possible, however, that there are differences in the specificity of the antibodies, especially as AChR is a high molecular-weight immunogen which presumably contains multiple potential immuno-potent determinants. The susceptibility to EAMG may thus be determined by a genetically controlled ability to respond to a specific determinant (or determinants), not necessarily identical with the immunogenic determinants. Alternatively, circulating anti-AChR may not have a major role in the pathogenesis of EAMG. The mouse model disease which has a great advantage because of the availability of many inbred strains, together with immunochemical analysis of the AChR molecule can be helpful in the elucidation of this question.

The humoral response against AChR is T-cell dependent as was shown in adoptive transfer experiments (Table 3). It seems that T cells are required both for a primary and secondary immune response. Lennon *et al.*[18] have previously demonstrated in rats that thymus-derived lymphocytes are required for the induction of EAMG and antibody to AChR. It should be interesting to try to dissociate the role of the thymus as a source of T-helping and T-suppressing cells from the possible specific role of the thymus as an antigenically cross reacting target with AChR[19].

We have demonstrated the induction of EAMG in mice after two injections of *Torpedo californica* AChR. The clinical, pharmacological, electrophysiological and immunological findings in mice with EAMG closely parallel those in patients suffering from MG as well as in other animals with EAMG. Mice with EAMG were shown to react with self-AChR, probably as a result of breakdown of tolerance. Strains seem to vary in their susceptibility to the experimental disease. We do not know, however, whether there is a similar pattern for susceptibility to spontaneous disease, and whether such a pattern is genetically controlled.

We thank Drs O. Abramsky, A. Aharonov and E. Mozes for discussions.

SARA FUCHS
DANIELA NEVO
REBECA TARRAB-HAZDAI

*Department of Chemical Immunology,
Weizmann Institute of Science,
Rehovot, Israel*

ISRAEL YAAR

*Department of Neurology,
Hadassah University Hospital,
Jerusalem, Israel*

1 Patrick, J., and Lindstrom, J., *Science*, **180**, 871–872 (1973).
2 Heilbronn, E., Mattsson, C., Stalberg, E., and Hilton-Brown, P., *J. neurol. Sci.*, **24**, 59–64 (1975).
3 Aharonov, A., Kalderon, N., Silman, I., and Fuchs, S. *Immunochemistry*, **12**, 765–771 (1975).
4 Green, D. P. L., Miledi, R., and Vincent, A., *Proc. R. Soc.*, **B189**, 57–68 (1975).
5 Tarrab-Hazdai, R., Aharonov, A., Abramsky, O., Yaar, I., and Fuchs, S., *J. exp. Med.*, **142**, 785–789 (1975).
6 Lennon, V. A., Lindstrom, J. M., and Seybold, M. E., *J. exp. Med.*, **141**, 1365–1375 (1975).
7 Tarrab-Hazdai, R., Aharonov, A., Silman, I., Fuchs, S., and Abramsky, O., *Nature*, **256**, 128–130 (1975).
8 Almon, R. R., Andrew, C. G., and Appel, S. H., *Science*, **186**, 55–57 (1974).
9 Aharonov, A., Abramsky, O., Tarrab-Hazdai, R., and Fuchs, S., *Lancet*, ii, 340–346 (1975).
10 Bender, A. N., Ringel, S. P., Engel, W. K., Daniels, M. P., and Vogel, Z., *Lancet*, i, 607–609 (1975).
11 Abramsky, O., Aharonov, A., Webb, C., and Fuchs, S., *Clin. exp. Immun.*, **19**, 11–16 (1975).
12 Simpson, J. A., *Scot. med. J.*, **5**, 419–436 (1960).
13 Lennon, V. A., and Carnegie, P. R., *Lancet*, i, 630–633 (1971).
14 Osserman, K. E., and Jenkins, G., *Mount Sinai J. Med.*, **38**, 497–537 (1971).
15 Oppenheim, J. J., Dougherty, S., Chan, S. P., and Baker, J., in *Laboratory Diagnosis of Immunological Disorders* (edit. by Vyas, G. N., Stiles, D. P., and Brecher, G.), 87–109 (Grune and Stratton, New York, 1975).
16 Lindstrom, J., and Lennon, V. A., *Ann. N.Y. Acad. Sci.* **274**, 254–274 (1976).
17 McDevitt, H. O., and Tyan, M. L., *J. exp. Med.*, **128**, 1–11 (1968).
18 Lennon, V. A., Lindstrom, J. M., and Seybold, M. F., *Ann. N.Y. Acad. Sci.*, **274**, 283–299 (1976).
19 Aharonov, A., Tarrab-Hazdai, R., Abramsky, O., and Fuchs, S., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 1456–1459 (1975).

Selective IgA deficiency in chickens with spontaneous autoimmune thyroiditis

SELECTIVE IgA deficiency is believed to be the commonest form of human immunodeficiency disease, and occurs with a frequency of between 1:500 and 1:700 in random populations[1,2]. Although the isolated deficiency is often considered non-pathological, there is clinical and laboratory evidence that IgA-deficient individuals suffer more sinopulmonary infection and autoimmune and gastrointestinal disorders[3,4]. Family studies have suggested a genetic factor but no defined inheritance patterns have been established[4,5]. The chicken model has been used for a number of years in our laboratory to study the ontogeny of immunoglobulin class development[6]. Surgical bursectomy at hatching[7,8], treatment of chick embryos with specific antisera[7,9-11] and combined neonatal bursectomy–thymectomy[12] can result in depressed levels of IgA. This report describes recent observations of a selective IgA-deficiency that spontaneously occurs in the Obese strain (OS) of White Leghorn chickens. The OS chickens develop spontaneous autoimmune thyroiditis (SAT) at several weeks of age[13]. Clinical hypothyroidism is

been reported that selective IgA deficiency in humans is often associated with increased amounts of serum IgM as well as increased number of plasma cells containing IgM in the laminia propria of the small intestine, indicating that a compensatory mechanism may occur in some individuals[3].

The relationship of autoimmune disease in humans to selective IgA deficiency has not been determined. While the majority of patients with autoimmune disorders have normal levels of IgA, those patients with the IgA deficiency have an increased incidence of abnormal κ/λ ratio, elevated levels of IgG and/or IgM, and serum 7S IgM[4]. Furthermore, those individuals who lack IgA have normal or increased numbers of IgA-bearing lymphocytes in the peripheral blood[20]. These observations suggest that the abnormality in the immune response occurs in the latter stages of cell maturation and results in the absence of IgA-secreting plasma cells. With the advent of an animal model for spontaneous IgA deficiency it is now possible to study the genetic and cellular events responsible for this deficiency. Indeed, preliminary experiments in our laboratory have already suggested a genetic relationship between IgA deficiency and alleles at the major histocompatibility locus in chickens.

Table 1 Serum immunoglobulin concentrations in OS and CS chickens*

Line	Age	Number tested	IgY(mg %)† ±s.d.	IgM(mg %) ±s.d.	IgA(mg %) ±s.d.	Number with undetectable IgA	Total number with <10 mg % IgA
OS	6 weeks	12	228±2.9	146±1.3	5.57±3.9	4	7
CS	6 weeks	18	254±2.0	93.4±1.3	24.5±1.3	0	0
OS	1 yr	30	456±1.7	201±1.5	16.5±4.4	5	8
CS	1 yr	24	599±1.3	89.3±2.2	27.6±2.1	0	1

*Minimum sensitivity of the radial immunodiffusion assay is 2.0 mg % for IgY and 3.0 mg % for IgM and IgA.
†Serum immunoglobulin levels are presented as geometric mean. Analyses were performed on $\log_{10}$ transformed data.

observed and high titres of circulating as well as bound anti-thyroglobulin antibodies can be detected[14]. The OS chickens were initially derived from the Cornell C-strain which reveals a 1% incidence of SAT. Selective breeding of CS birds, which showed phenotypic symptoms of hypothyroidism, resulted in the OS of which now, in generations 16 and 17, show 96% SAT. The histological and serological features in OS chickens make it a reasonable animal model to study Hashimoto's thyroiditis, an autoimmune disorder in humans[15].

Serum immunoglobulins IgY (ref. 16), IgM, and IgA were quantified by radial immunodiffusion in Agarose gel with heavy chain-specific antisera and reference standards prepared in our laboratory[17]. This technique is conventionally used to diagnose selective IgA deficiency in humans[18]. Both systems can detect IgA concentrations as low as 2–3 mg %.

In Table 1 the serum immunoglobulin concentrations of a random population of 6-week and 1-yr-old OS and CS chickens are presented. Serum IgA concentrations were markedly decreased in OS chickens with a mean concentration of 5.6 mg % in 6-week-old birds and 16.5 mg % in 1-yr-old birds. Although many of these OS birds had normal levels of IgA, 33% (4 of 12) 6-week-old birds and 16% (5 of 30) 1-yr-old birds had undetectable levels of IgA (<3.0 mg %). In contrast, only one CS bird had an IgA level of less than 10 mg % and none had undetectable levels of IgA. In addition to the isolated IgA deficiency in OS chickens, an increase in serum IgM levels were noted. It was observed, however, that the IgA-deficient birds did not necessarily have a selective increase in IgM, but rather the increased IgM was present throughout the entire OS population. This is consistent with the increased number of IgM-bearing B-cells found in OS chickens[19]. It has previously

OS strain was developed and is maintained by the Department of Poultry Science at Cornell University. Dr Luster is a recipient of a USPHS fellowship from the NIAID.

MICHAEL I. LUSTER
GERRIE A. LESLIE

Department of Microbiology and Immunology,
University of Oregon Health Sciences Center,
Portland, Oregon 97201

RANDALL K. COLE

Department of Poultry Science,
New York State College of Agriculture and Life Sciences,
Cornell University,
Ithaca, New York 14853

[1] Bachmann, R., *Scand. J. clin. lab. Invest.*, **17**, 316–320 (1965).
[2] Hobbs, J. R., *Lancet*, i, 110–114 (1968).
[3] Ammann, A. J., and Hong, R., *Medicine*, **50**, 223–236 (1971).
[4] Horowitz, S., and Hong, R., in *Immunodeficiency in Man and Animals*, 11 (edit. by Bergsma, D.), 129–133 (Sinauer Associates, Massachusetts, 1975).
[5] Grundbacher, F. J., *J. med. Genet.*, **9**, 344–347 (1972).
[6] Leslie, G. A., *Am. J. Vet. Res.*, **36**, 482–485 (1975).
[7] Kincade, P. W., and Cooper, M. D., *Science*, **179**, 398 (1973).
[8] Martin, L. N., and Leslie, G. A., *Proc. Soc. exp. Biol. Med.*, **143**, 241–243 (1973).
[9] Martin, L. N., and Leslie, G. A., *J. Immun.*, **113**, 120–126 (1974).
[10] Leslie, G. A., and Martin, L. N., *Int. Arch. Allergy*, **46**, 834–841 (1974).
[11] Leslie, G. A., Stankus, R. P., and Martin, L. N., *Int. Arch. Allergy*, **51**, 175–185 (1976).
[12] Perey, D. Y. E., and Bienenstock, J., *J. Immun.*, **111**, 633–637 (1973).
[13] Witebsky, E., Kite, J. H., Wick, G., and Cole, R. K., *J. Immun.*, **103**, 708–715 (1969).
[14] Cole, R. K., *Genetics*, **53**, 1021–1033 (1966).
[15] Wick, G., Sundick, R. S., and Albini, B., *Clin. Immunopath.*, **3**, 272–300 (1974).
[16] Leslie, G. A., and Clem, L. W., *J. exp. Med.*, **130**, 1337–1352 (1969).
[17] Leslie, G. A., and Martin, L. N., *J. Immun.*, **110**, 1–9 (1973).
[18] Buckley, R. H., in *Immunodeficiency in Man and Animals*, 11 (edit. by Bergsma, D.), 134–142 (Sinauer Associates, Massachusetts, 1975).
[19] Albini, B., and Wick, G., *Nature new Biol.*, **249**, 653–654 (1974).
[20] Lawton, A. R., Royal, S. A., Self, S., and Cooper, M. D., *J. lab. clin. Med.*, **80**, 26–33 (1972).

Transformation of human lymphocytes by Epstein-Barr virus is inhibited by phosphonoacetic acid

PHOSPHONOACETIC acid (PAA) is an inhibitor of herpes simplex virus and of several other herpes viruses. It acts as a specific inhibitor of two viral DNA polymerases, those of herpes simplex and of Marek's disease (a herpes-induced leukaemia in chickens)[1,2]. We have used PAA as a probe to investigate the possible existence of a virus-directed DNA polymerase and its role in replication of viral genomes during the transformation and proliferation of normal human peripheral lymphocytes induced by Epstein-Barr virus (EBV), which is a herpes virus closely associated with the development of several neoplasms in man.

Synthesis of the EBV capsid antigen (VCA), but not of the early antigen (EA), is dependent on DNA synthesis[3]. Moreover, EBV producer lines cease synthesis of VCA and biologically active virus after several days of treatment with PAA at concentrations which do not have a measurable effect on cellular growth, cellular DNA synthesis or expression of either EA or the EBV nuclear antigen (EBNA)[4,5]. The average number of copies of the viral genome per cell also decreases in these conditions because of the inhibition of productive viral DNA synthesis. Replication of latent viral DNA seems to be unaffected by PAA[5,6]. Although no EBV-specific DNA polymerase has been isolated, these data suggest the existence of such an enzyme.

Little is known about the events by which EBV causes resting normal peripheral lymphocytes to transform and proliferate. The only two indicators of transformation are an increase in cellular DNA synthesis[7,8] and the appearance of EBNA several days after infection[9]. Because of the limitations of material and available techniques, information about the behaviour of the viral DNA during infection is limited. Replication of the viral DNA after infection is suggested, however, by the presence of multiple copies of the viral genome in lymphoblastoid cell lines produced by *in vitro* infection with EBV[10].

EBV transforms only the B-cell population of peripheral blood lymphocytes. Therefore, purified B-cell populations from normal adults were used to study DNA synthesis after infection with EBV. Use of purified B cells greatly decreases the background level of DNA synthesis by untransformed cells and provides a much more sensitive measure of virus-induced DNA synthesis than does the use of unfractionated cells. This method is at least as sensitive as the more commonly used cord blood lymphocytes and a full account will be published elsewhere. Briefly, blood was drawn from normal healthy adult donors into heparinised syringes and mixed 1 : 1 with Hanks' balanced salt solution. The peripheral lymphocytes were purified by centrifuging over Ficoll-Hypaque[11]. The lymphocytes at the interface were washed with medium (RPMI 1640 plus penicillin (100 U ml⁻¹), streptomycin (100 µg ml⁻¹), glutamine (300 µg ml⁻¹) and 5% foetal calf serum) and passed over an immunoabsorbent column of goat (anti-human Fab') antibody coupled to Sephadex G-200 (ref. 12). The B-cell population was bound to the column and was eluted with a 1% solution of human γ globulin and then twice washed with medium. This population is > 95% B cells as judged by labelling with fluorescent rabbit (anti-human Fab') antibody.

The rate of DNA synthesis by normal peripheral B lymphocytes cultured in the presence of various concentrations of PAA and assayed 6–7 d after addition of the transforming strain of EBV from the supernatant of the B95-8 marmoset lymphoblast culture is shown in Fig. 1.

There was demonstrable inhibition of DNA synthesis with as little as 12.5 µg ml⁻¹, with 50% inhibition at 100 µg ml⁻¹ and complete inhibition at 200 µg ml⁻¹. As a control, the effect of PAA on DNA synthesis in B lymphocytes stimulated by phytohaemagglutinin or pokeweed mitogen was studied. There was no inhibition up to a PAA concentration of 100 µg ml⁻¹ and only 30% inhibition at 200 µg ml⁻¹.

Consistent with the above observations of DNA synthesis, PAA at 100–200 µg ml⁻¹ was sufficient to inhibit completely the outgrowth of B-cell cultures infected with the B95-8 strain and observed for up to 1 month after infection, compared with controls which grew out after 12 d (Fig. 2).

The inhibitor did not act directly on the virus, however,
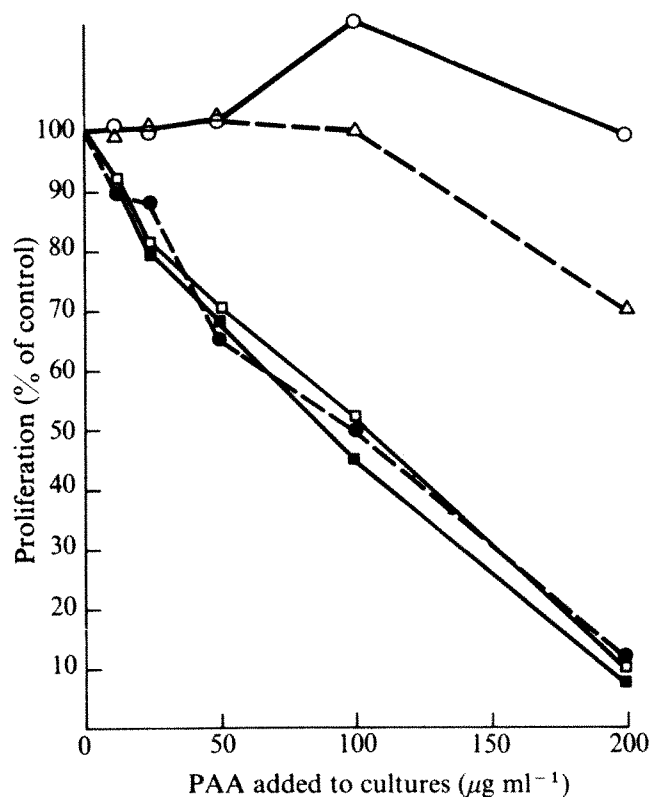
Fig. 1 Effect of PAA on DNA synthesis of normal peripheral B lymphocytes transformed by EBV or mitogens. Purified B cells were suspended at 2×10^6 ml⁻¹ in RPMI 1640 medium supplemented with 20% foetal calf serum, and 0.1-ml samples were added to Wells of a Linbro Micro Test II plate. EBV (see below), PHA (Burroughs Wellcome final concentration 0.25 µg ml⁻¹) and pokeweed mitogen (Gibco, final concentration 40 µg ml⁻¹) were diluted appropriately with the same medium and added in 0.1 ml to make a final volume of 0.2 ml per well. PAA, at appropriate dilutions, was added when required in a volume of 3–5 µl by means of an automatic delivery Hamilton syringe. All experiments were set up in triplicate. The cells were incubated at 37 °C in a humidified, 5% CO_2 atmosphere. After 6 d the cells were pulsed overnight with 50 µl of medium containing ³H-thymidine (New England Nuclear, 2 Ci mmol⁻¹, 2 µCi ml⁻¹). The cells were collected on glass fibre filters, washed repeatedly with saline followed by 5% ice-cold trichloroacetic acid and 95% ice-cold ethanol. The filters were dried and the remaining radioactivity was assessed by addition of 5 ml of Liquifluor scintillation fluid and counting. The experimental results are expressed as

$$\frac{\text{c.p.m. in experimental}}{\text{c.p.m. in positive control}} \times 100\%$$

The c.p.m. in the positive controls (EBV or mitogen-added but no PAA) were EBV 1:10 (●)=6,810; 1:50 (□)=6.490; 1:100 (■)=1,510; PHA (○)=6,764; pokeweed mitogen (△)=4,774. Broken and unbroken lines represent B cells from different individuals with unstimulated incorporations of 105 and 141 c.p.m. respectively. The B95-8 EBV concentrates used in these experiments were provided by Dr K. A. Traul of Pfizer Inc. lot no. 3373-11, TCID₅₀ 10⁶·⁵ U ml⁻¹ by induction of EBNA in cord blood, electron microscopic count 4×10^8 viral particles per ml.
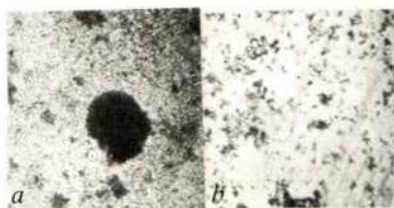
Fig. 2 Appearance of cultures infected with EBV to which PAA had not (*a*) or had (*b*) been added. Parallel cultures were infected with a supernatant from the B95-8 marmoset cell line, at a dilution of 1:10. Every 3–4 d half the culture medium was replaced with fresh medium. For full experimental details, see legend to Fig. 1a, Control culture infected with EBV alone and photographed 12–15 d after infection. After this time it is necessary to transfer the cells to larger wells to maintain them. Growth, as illustrated, occurred in all of 33 wells to which virus was added. *b*, Culture infected with EBV and maintained in the presence of PAA 100 μg ml⁻¹ photographed 30 d after infection. At no time was an increase in cell number observed in any of 33 wells.

for preincubation with PAA for up to 6 h at 37 °C did not affect the activity of the virus in stimulating DNA synthesis after removal of the inhibitor (data not shown). Furthermore, expression of at least two viral functions was unaffected insofar as PAA up to 200 μg ml⁻¹ did not affect induction of EBNA in the peripheral B lymphocytes, nor did it affect induction of EA in the Raji cell line by the superinfecting strain of EBV from the P3HR-1 cell line (Table 1). Although PAA does not inhibit induction of EBNA in either adult peripheral B lymphocytes or those from cord blood, it is inhibited by cytosine arabinoside which blocks all DNA synthesis (I. Ernberg and G. Klein, personal communication). Both PAA and cytosine arabinoside are ineffective, however, in preventing EBNA induction during conversion of the EBV-negative lymphoblastoid line Ramos

into its EBV-positive derivatives[13]. This suggests that cytosine arabinoside does not act directly on EBNA production but is effective in the transformation experiments because it prevents blastogenesis, and that once the cells have undergone blastogenesis, EBNA may be induced without DNA synthesis.

It is interesting that with PAA at 100–200 μg ml⁻¹, cellular DNA synthesis stimulated by mitogens and EBNA induction were unaffected, whereas the transformation and outgrowth of EBV-infected cultures was inhibited. This implies that induction of EBNA is not sufficient to cause
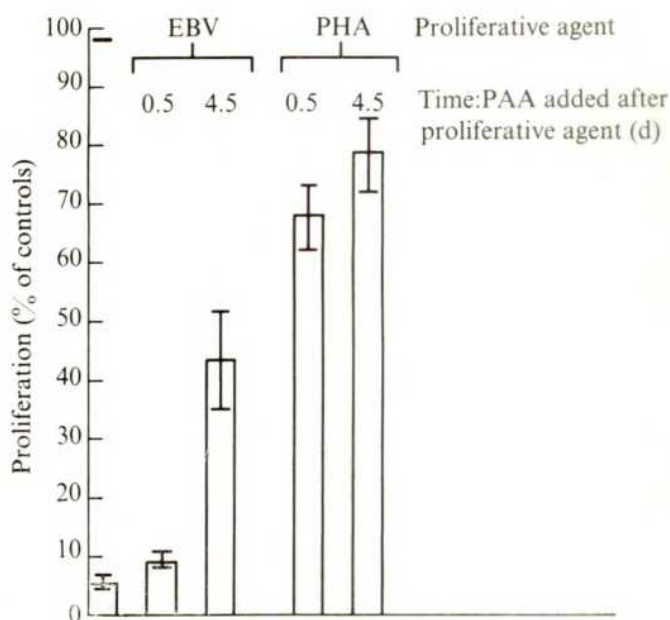
Fig. 4 Effect of adding PAA (200 μg ml⁻¹) 0.5 and 4.5 d after stimulation of DNA synthesis by EBV or PHA. Normal peripheral B lymphocytes were treated with EBV (dilution 1:10) or PHA as described in the legend to Fig. 1. PAA was added in 3–5 μl either 0.5 d or 4.5 d after cultures were set up. The rate of DNA synthesis was assayed 6 d after cultures were infected. Each column is the result of averaging the results obtained using cells from six different donors and the bars represent 1 s.d. The positive controls varied between 700 and 3,700 c.p.m. with EBV and 2,200 to 2,600 with PHA. A culture not treated with virus is shown in the left column.

Fig. 3 Effect of PAA, added at various times after infection, on the rate of EBV-stimulated DNA synthesis. Purified human peripheral B lymphocytes were set up in culture in the presence of a 1:10 dilution of EBV as described in the legend to Fig. 1. PAA was added at various times in a volume of 3–5 μl to give the appropriate final concentration from an automatic delivery Hamilton syringe. The cells were assayed for DNA synthesis, as described in the legend to Fig. 1, 9 d after infection. ●, PAA final concentration 50 μg ml⁻¹; ○, PAA concentration 200 μg ml⁻¹. The c.p.m. in positive control = 1,976.

cells to proliferate logarithmically even when the cellular DNA synthetic system is intact.

The nature of inhibition of DNA synthesis by PAA after EBV infection was studied further by adding the inhibitor to different cultures at various times after the virus. The rate of DNA synthesis was measured 9 d after infection. PAA was effective when added up to 3 d after infection but thereafter effectiveness decreased (Fig. 3). The effectiveness of PAA when added many hours after the virus confirms the observation that PAA does not act on the viral particle directly and implies also that it does not affect penetration of the cell by the virus. The decrease in effectiveness of PAA could have been a result of its slow incorporation; it took about 5 d to reach effective levels. This was ruled out, however, by the observation that the decrease in sensitivity occurred at the same time whether the cells were collected for assay 7 or 9 d after infection.

The decrease in the sensitivity of EBV-stimulated DNA synthesis to PAA is specific, as shown when PAA at 200 μg ml⁻¹ was added to cultures 0.5 or 4.5 d after infection with EBV or stimulation with PHA (Fig. 4). There was little change in the sensitivity of PHA-stimulated cellular DNA synthesis with various times of addition, synthesis being inhibited by 20–30% in both cases, perhaps because of
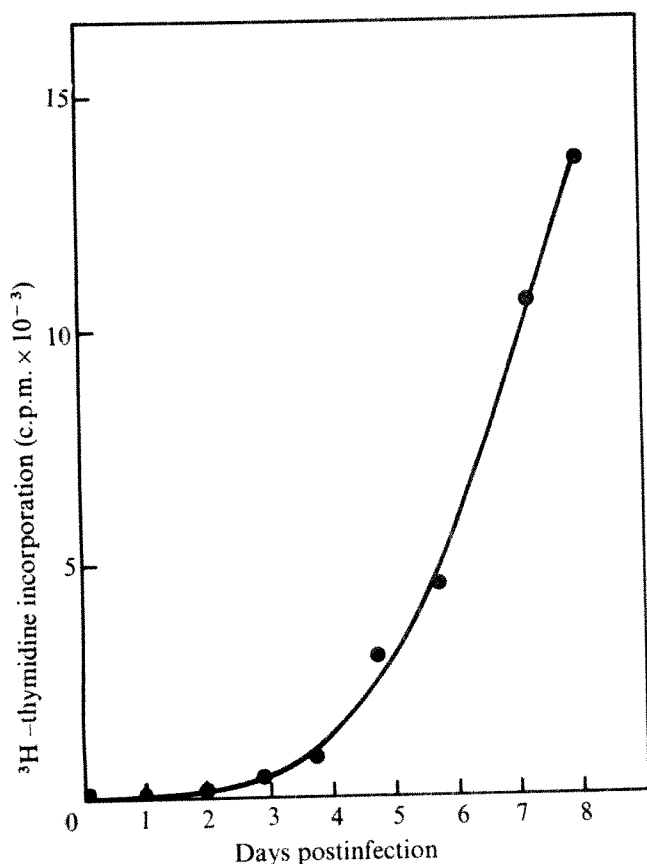
Fig. 5 Time course of DNA synthesis in lymphocytes infected with EBV. Parallel cultures of uninfected and virus-infected (by additions of supernatant solution from a culture of B95-8 cells diluted 1:10) cells were set up in triplicate at various times. Thereafter, the rate of DNA synthesis was assayed as before, except the cells were pulsed for 4 h. The specific rate of DNA synthesis was calculated by subtracting the value measured for the uninfected cultures from that obtained with the infected cultures. For full experimental details, see legend to Fig. 1.

Table 1 Effect of PAA on induction of EBNA in normal peripheral B lymphocytes and EA in the Raji lymphoblastoid cell line

| | | (PAA) μg ml^{-1} | | |
	no virus	0	100	200
EBNA+ve cells	—	+	ND	+
% EA+ve cells*	<0.25	5.1	5.8	4.8

To measure EBNA induction, 5×10^6–10×10^6 B lymphocytes were infected with B95-8 virus in the usual way (see legend to Fig. 1); 3–4 d later the cells were collected and stained for EBNA by the anticomplement immunofluorescent technique[14]. EA induction by P3HR-1 EBV in the Raji cell line was tested as described by Klein *et al.*[15]. Anti-EA antisera for the direct fluorescence test were provided by Dr G. Klein.
*Average of duplicates, about 2,000 cells each were counted.
ND, Not done; +, the same level of positive cells as in the control, approximately 5%.

is then kept constant—it replicates with the cellular DNA. The amplification may conceivably be a prerequisite for integration of one or more EBV genomes into the cellular DNA synthetic system, functionally and perhaps also physically. Alternatively, replication could be required simply for the transcription and translation of one or more genes required for transformation.

If a PAA-sensitive, virally induced DNA polymerase (by analogy with herpes simplex[1] and the Marek's disease[2] DNA polymerase) is required for gene amplification before integration, then it follows that viral DNA synthesis is required for true cellular proliferation. At the simplest level, the single cell infected with EBV in the presence of PAA may be transformed, able to express EBNA and to divide indefinitely. It carries only one non-replicating copy of the viral genome, however, and therefore can generate only one daughter cell carrying and expressing the viral genome and able to divide. Thus, cultures infected with EBV in the presence of PAA may contain transformed and dividing cells, but the cell number per culture would not increase exponentially and inhibition of proliferation would be observed.

DAVID THORLEY-LAWSON
JACK L. STROMINGER

Sidney Farber Cancer Institute,
Harvard Medical School,
Boston, Massachusetts 02115

1 Mao, J. C-H., and Robishaw, E. E., *Biochemistry*, 14, 5475–5479 (1975).
2 Leinbach, S. S., Reno, J. M., Lee, L. F., Isbell, A. F., and Boezi, J. A., *Biochemistry*, 15, 426–429 (1976).
3 Gergely, L., Klein, G., and Ernberg, I., *Int. J. Cancer*, 7, 293–302 (1971).
4 Nyormoi, O., Thorley-Lawson, D. A., Elkington, J., and Strominger, J. L., *Proc. natn. Acad. Sci. U.S.A.*, 73, 1745–1748 (1976).
5 Summers, W. C., and Klein, G., *J. Virol.*, 18, 151–155 (1976).
6 Yajima, Y., Tanaka, A., and Nonoyama, M., *Virology*, 71, 352–354.
7 Gerber, P., and Hoyer, B. H., *Nature*, 231, 46–47 (1971).
8 Robinson, J., and Miller, G., *J. Virol.*, 15, 1065–1072 (1975).
9 Moss, D. J., and Pope, J. H., *Int. J. Cancer*, 15, 503–511 (1975).
10 Pagano, J. S., *J. infect. Dis.*, 132, 204–223 (1975).
11 Boyum, A., *Scand. J. clin. Lab. Invest.*, 21, 31–89 (1968).
12 Chess, L., MacDermott, R. P., Sondel, P. M., and Schlossman, S. F., *Prog. Immun. II*, 3 (edit. by Brent, L., and Holborrow, J.),(North-Holland, Amsterdam, 1974).
13 Klein, G., Giovanella, B., Westman, A., Stehlin, J. S., and Mumford, D., *Intervirology* (in the press).
14 Reedman, B. M., and Klein, G., *Int. J. Cancer*, 11, 499–520 (1973).
15 Klein, G., Dombos, L., and Gothoskar, B., *Int. J. Cancer*, 10, 44–57 (1972).

inhibition of a host DNA polymerase which is reported to be sensitive to PAA (J. Liversidge and W. Benz, personal communication). The EBV-stimulated DNA synthesis was completely inhibited when PAA was added early after infection but only about 50% inhibited when added at 4.5 d.

The most likely explanation of these results is that there is a process which is sensitive to PAA and is required for the initiation of cellular proliferation. This process occurs during the first 3–4 d after infection, at which time cellular DNA synthesis begins and apparent PAA sensitivity decreases. This hypothesis is supported by the observation that a considerable increase in cellular DNA synthesis occurs at 3–4 d after EBV infection (Fig. 5), precisely the time after which addition of PAA is no longer effective in inhibiting transformation.

In conclusion, the data presented here indicate the existence of a PAA-sensitive event, which is necessary for transformation and unlimited proliferation of normal peripheral B lymphocytes. PAA does not, however, act directly on the virion nor does it block penetration of the cell by the virus or the subsequent expression of EBNA. It is known that cell lines produced by *in vitro* infection with EBV contain 10–15 copies of the viral genome per cell[10]. It seems likely therefore that some amplification of the viral genome occurs after infection. This amplification must cease after a short time and the viral genome number

Complement-dependent adherence of mast cells to schistosomula

MAST cells are thought to have a major role in the host response to parasitic helminths. Infection with these organisms is often accompanied by anaphylactic hypersensitivity and mast cell infiltration[1]. In several experimental models, host resistance to helminth infection has been shown to be dramatically suppressed as a result of mast cell depletion[1]. For example, after injection with reserpine, laboratory mice are unable to mount a cutaneous inflammatory response against *Schistosoma mansoni* (A.S.,

P. W. Askenase, S. McIntyre and F. von Lichtenberg, un-published) and after pretreatment with compound 48/80 can no longer immunologically reject challenge infections of this parasite[2]. Interactions between mast cells and helminths have generally been attributed to soluble antigens released into tissue fluids. Described here is a series of *in vitro* experiments with schistosomula of *S. mansoni* which demonstrates that mast cells are also capable of reacting directly with parasite surfaces. This interaction appears to depend on the recognition of parasite-bound complement by receptors on the mast-cell membrane.

Schistosomula were prepared by allowing *S. mansoni* cercariae to penetrate rat skin *in vitro*[3], and were stored overnight at 4 °C before use[4]. Mast cells were purified from peritoneal washouts of uninfected male Charles River CD (Caesarean derived) rats (175–300 g) by centrifugation over Metrizamide (22.5%)[5]. The cells recovered from the pellet fraction were routinely found to contain a minimum of 90% mast cells as determined by toluidine-blue staining. Both schistosomula and mast cells were washed three times in HEPES-buffered Hanks' minimal essential medium (MEM) containing 1% heat-inactivated foetal calf serum (FCS) and resuspended in MEM containing 10% FCS (MEM/FCS). A serum pool obtained from CD rats 8 weeks after a primary infection[6] of 1,500 cercariae each was used as a source of anti-schistosome antibody. The immune serum was diluted in MEM/FCS and used in the assay at a final concentration of 1/20. The adherence test was carried out
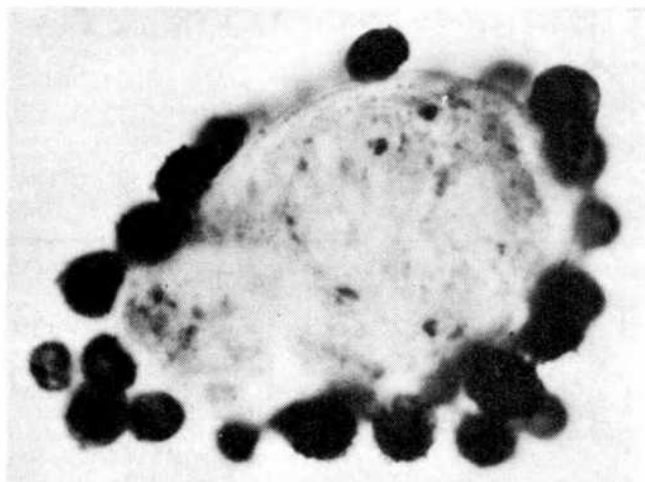
Fig. 1 Photomicrograph ($\times 216$ magnification) of a typical schistosomulum after incubation with fresh immune rat serum and purified rat mast cells.

in glass tubes (6×50 mm) containing the following sequential additions: 0.1 ml of schistosomula (90–110 organisms), 0.1 ml of diluted immune serum, 0.1 ml of diluted normal rat serum (complement source), and 0.1 of mast cells (2×10^5–4×10^5 total cells). After incubation at 37 °C for 90 min, the supernatant was withdrawn from each tube and the cell–parasite sediment resuspended in 0.1 ml of MEM/FCS containing 0.01% toluidine-blue. The parasites were then examined microscopically at 80 times magnification for adhering mast cells. Organisms with more than five adherent cells were scored as positive. Each result is the average of readings on replicate assay tubes scored independently by two different observers.

Schistosomula incubated with fresh immune rat serum formed 'rosettes' with the purified mast cells (Fig. 1), each positive organisms having a coat of about 20–100 adherent, toluidine-blue-positive cells. Approximately 90% (Table 1) of the larvae in the incubation mixture reacted in this fashion. In contrast, few if any mast cells adhered to schistosomula incubated in heat-inactivated (2 h, 56 °C) immune serum (Table 1). The addition of fresh normal rat serum to the heat-treated immune serum restored its full activity, an observation which suggested that the heat-labile factor was complement

Table 1 Adherence of rat mast cells to schistosomula in the presence of rat sera

Immune rat serum	Normal rat serum	Mast cell coated organisms (%)
Fresh	None	89
Inactive	None	14
Inactive	Fresh (1/20)	93
Inactive	Fresh (1/40)	80
Inactive	Fresh (1/80)	33
Inactive	1 h, 56 °C (1/20)	21
Inactive	30 min, 50 °C (1/20)	32
Inactive	Fresh (1/20)+EDTA (5×10^{-3} M)	3
Inactive	Fresh (1/20)+EGTA (5×10^{-3} M)	89
Inactive	Fresh (1/20)+Cobra factor	35
Inactive	+Zymosan, 1 h, 37 °C (1/20)	8
Inactive	+Zymosan, 2 h, 17 °C (1/20)	20
None	Fresh (1/20)—Pool A	46
None	Fresh (1/20)—Pool B	85
None	*Preincubation in Pool B (1/20)	85
None	*Preincubation in Pool B (1/20)+EDTA (5×10^{-3} M)	8
None	None	1

*Schistosomula preincubated with rat serum (1 h, 37 °C) and washed three times (with MEM/FCS) before addition of mast cells.

rather than antibody. This hypothesis was supported by a series of experiments (Table 1) in which the adherence-promoting activity of fresh normal rat serum was shown to be partially or completely abolished by the following treatments known to inactivate complement[7,8]: (1) heating for 1 h at 56 °C; (2) pre-incubation with either zymosan (15 mg ml^{-1} serum) or cobra venom factor (200 U ml^{-1} serum) for 1 h, at 37 °C; (3) the addition of EDTA (5×10^{-3} M) to the cultures. The dependence of the adherence reaction on the alternate pathway of complement[9,10] was suggested by the failure of EGTA (5×10^{-3} M)[11] to inhibit rosetting and by the reduction in activity observed after heating the normal serum at 50 °C for 30 min (ref. 9) or pretreating it with zymosan at low temperature (2 h, 17 °C)[9].

Schistosomula incubated in normal rat serum without immune serum as an antibody source were also found to bind mast
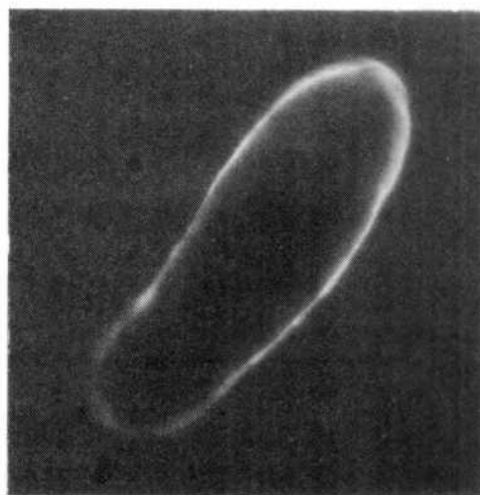
Fig. 2 Photomicrograph ($\times 224$) of a typical schistosomulum, preincubated with normal rat serum (Pool B), washed three times in MEM/FCS and stained with fluorescein-labelled rabbit antibody specifiic for rat C3 (from Cappel Laboratories). No evidence of staining was observed when the same organisms were treated with fluorescein-labelled rabbit antibody (Cappel) produced against rat IgG but cross reactive with rat immunoglobulins of other classes. In contrast, schistosomula preincubated with immune rat serum stained brightly (as above) with the same anti-immunoglobulin reagent.

cells, different serum pools inducing different levels of adherence (Table 1). After preincubation with normal rat serum (1 h, 37 °C) and extensive washing (with MEM/FCS), schistosomula remained highly reactive with mast cells. The activity of normal serum in promoting adherence did not seem to be due to the presence of 'natural antibodies' with specificity for schistosome surface antigens since membrane-bound rat antibodies were not detected on worms pretreated with normal serum by means of immunofluorescent staining with fluorescein-labelled antibody specific for rat immunoglobulins. In contrast the same pretreated organisms stained brightly (Fig. 2) with an immunofluorescent reagent specific for rat C3. Neither C3 staining nor mast-cell adherence was observed if EDTA $(5 \times 10^{-3}$ M) was added to the normal rat serum in the preincubation mixture.

These results suggest that the adherence of mast cells to schistosomula is dependent on the recognition of parasite-bound complement by receptors on the mast cell surface. Complement in a form suitable for recognition by mast cells is acquired by the worms either as a consequence of the reaction of antibody at the parasite's surface or, in the absence of antibody, by direct fixation. Previous studies have indicated that cercariae of *S. mansoni* possess an alternative pathway-dependent complement (C3)-activating system[12]. The retention by schitosomsula of a similar complement-activating mechanism after conversion from the cercarial stage of the life cycle would explain their behaviour in the experiments reported here.

An important question raised by the mast cell–schistosomulum adherence phenomenon concerns the nature of the receptors on the mast-cell surface involved in the recognition of the parasite–complement complex. Rat mast cells have been shown to respond to the anaphylatoxins, C3a and C5a[13]. Since both complement fragments are normally released in the fluid phase and are extremely labile[13], it is unlikely that the adherence reaction involves receptors for these molecules. An alternative hypothesis is that mast cells, in common with platelets, primate erythrocytes, neutrophils, monocytes, macrophages and B lymphocytes, possess receptors for surface-bound activated C3 (refs 14, 15) and that the mast cell–schistosomulum interaction is dependent on this previously unrecognised set of immune adherence receptors.

Other questions raised by the adherence phenomenon concern the effects of the interaction on the release of mediators from mast cells and on the viability of the target organisms. Experiments aimed at answering these questions are now in progress.

I thank Susan McIntyre for technical assistance. This work was supported by grants from the Edna McConnell Clark and Rockefeller Foundations.

A. SHER

Department of Pathology,
Peter Bent Brigham Hospital and Harvard Medical School,
Boston, Massachusetts 02115

1 Ogilvie, B. M., in *Progress in Immunology II*, 4 (edit. by Brent, L. and Holborrow J.), 127–135 (North-Holland, Amsterdam, 1974).
2 Dean, D. A., Murrell, K. D., Minard, P., and Vannier, W. E., *Fedn Proc.* 35, 228 (abstr. No. 116) (1976).
3 Clegg, J. A., and Smithers, S. R., *Int. J. Parasit.*, 2, 79–98 (1972).
4 Butterworth, A. E., Sturrock, R. F., Houba, V., and Rees, P. H., *Nature*, 252, 503–505 (1974).
5 Paterson, N. A., Leid, R. W., Said, J. S., Wasserman, S. I., and Austen, K. F., in *Lung Cells in Disease* (edit. by Bouhuys, A.), (North-Holland, Amsterdam, in the press).
6 Smithers, S. R., and Terry, R. J., *Parisitology*, 55, 695–700 (1965).
7 Mayer, M., in *Experimental Immunochemistry* (edit. by Kabat, E. A.). 2nd ed. (Thomas, Springfield, Illinois, 1961).
8 Cochrane, C. G., Müller-Eberhard, H. J., and Aiken, B. S., *J. Immun.*, 105, 55–69 (1970).
9 Pillemer, L., *et al.*, *Science*, 120, 279–285 (1954).
10 Götze, O., and Müller-Eberhard, H. J., *J. exp. Med.*, 134, 905–1085 (1971).
11 Fine, D. P., Marney, S. R., Colley, D. G., Sergent, J. S., and Des Prez, R. M., *J. Immun.*, 109, 807–809 (1972).
12 Machado, A. J., Gazzinelli, G., Pellegrino, J., and Dias da Silva, W., *Expl Parasit.*, 38, 20–29 (1975).
13 Müller-Eberhard, H. J., and Vallota, E. H., in *Biochemistry of the Acute Allergic Reactions* (edit. by Austen, K. F. and Becker, E. L.), (Blackwell, Oxford, 1971).
14 Nelson, D. S., in *Ciba Foundation Symposium on Complement* (Churchill, London, 1965).
15 Lay, W. H., and Nuzzenzweig, V., *J. exp. Med.*, 128, 991–1007 (1968).

Variable Ca sensitivity of a K-selective channel in intact red-cell membranes

THE membranes of many cells exhibit a K-selective permeability mechanism with a gating process controlled by the concentration of Ca^{2+} in the cytosol[1]. Originally described in human red cells[2], where its function still remains unknown, this or similar mechanisms[3] seem to mediate important physiological actions in other cells[4–8].

The human red blood cell remains, nevertheless, the best model for the study of the interaction between Ca and the K gating mechanism because the composition of the media on both sides of the membrane can be known and controlled better than in any other cell[10–13]. This control is achieved, however, by treatments which are known to alter considerably the Ca reactivity of Ca-transporting mechanisms[14–16]. We
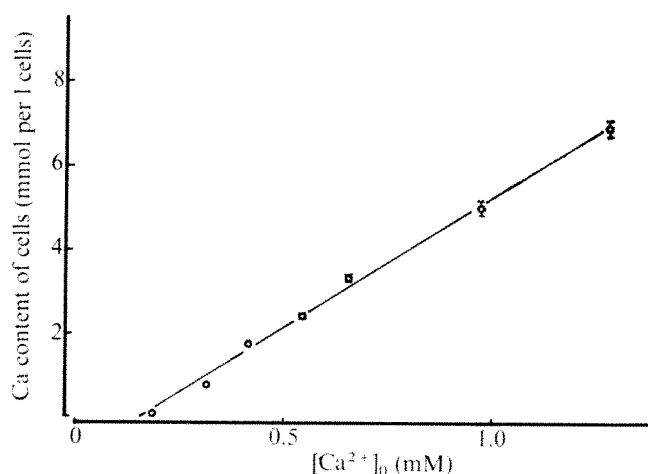
Fig. 1 Steady-state Ca content of intact red cells as a function of the external Ca^{2+} concentration in the presence of the divalent cation ionophore A23187. Red cells from fresh blood or from 3–4-d-old bank blood were washed as described previously[18] and then resuspended at about 10% haematocrit in a medium containing KCl, 75 mM; Tris-HCl (pH 7.5 at 37 °C), 10 mM; NaCl, 75 mM; $MgCl_2$, 1 mM; Inosine, 10 mM; 5 μCi of ^{45}Ca, and various concentrations of $CaCl_2$ up to about 2.5 mM. The total volume of suspension was 2.25 ml. At $t = 0$, 10 μl of absolute ethanol containing 0.05 to 1 mg ml^{-1} of the ionophore (0.1 mg ml^{-1} in this experiment) were added to the suspension which was in a magnetically stirred polythene vial inside a waterbath at 37 °C. At $t = 15.5$ min, 0.1 ml of isotonic ^{42}K–KCl was added to the suspension. At 15, 16, 16.5, 17, 18, 20, 25 and 30 min, 0.1 ml samples were placed in 1.5-ml Eppendorf centrifuge tubes containing 0.4 ml of n-butylphthalate (BDH, density 1.042–1.045) and 0.9 ml of an isotonic [Na + K]-medium containing 5 mM Tris-EGTA, all at 0–1 °C. Each sample was centrifuged immediately for 10 s at 12,000g in an Eppendorf centrifuge model 3200. The cells formed a compact pellet at the bottom, retaining only about 0.1% of the original extracellular fluid as estimated using ^{45}Ca in the presence of a large excess of EGTA. The aqueous supernatant remaining on top of the oil layer was removed by aspiration and the walls of the tube cleaned with cotton swabs. The cells were lysed and the proteins precipitated with 5% trichloroacetic acid (TCA) all in the same tube. Aliquots of the TCA-supernatant were used for counting both isotopes in the same samples. ^{45}Ca was counted after ^{42}K decay. Samples for total activity and haemoglobin were taken at the end of each run. In the media used here, the haematocrit did not change during the experiments and the internal K concentration remained constant over a thousandfold change in K permeability at about 130 mmol per l cell water. The results concerning the ^{42}K fluxes are given in Fig. 2. The Ca content of the cells was calculated as the mean ±2 s.e.m. of the seven time points. The slope of this particular curve is 5.91. The fraction of ionised Ca inside the cells, as calculated from this slope according to Ferreira and Lew[18], was 0.33. This value was used to estimate the concentration of ionised Ca in the cytoplasm of the cells in the experiment of Fig. 3 and an identical procedure was applied in the experiment of Fig. 4.
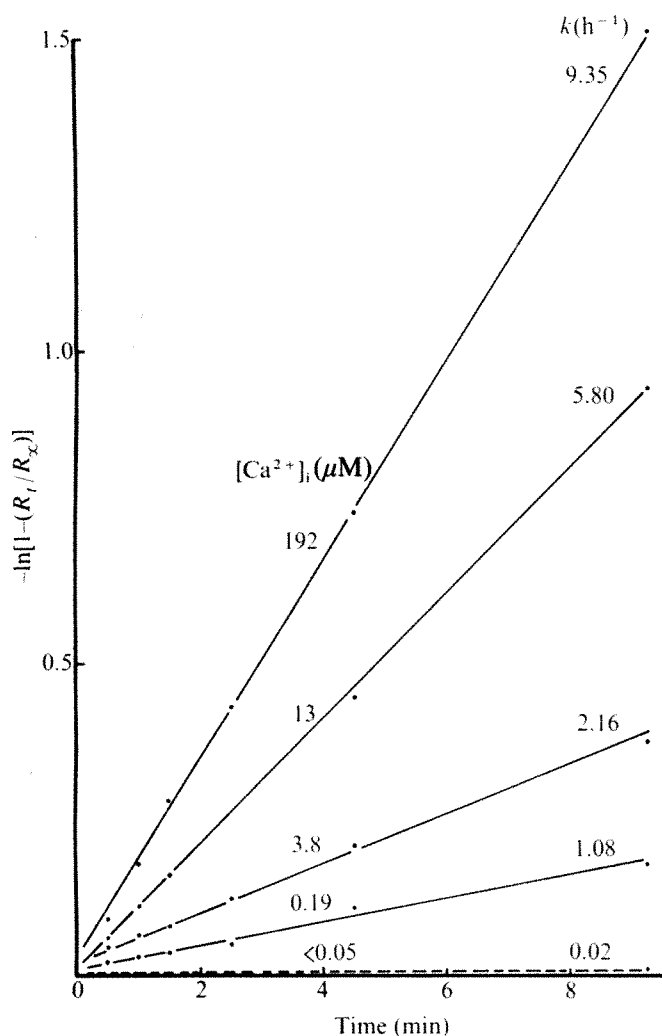
Fig. 2 The rate of tracer (^{42}K) equilibration in intact red cells at different intracellular levels of Ca^{2+}. The tracer equilibration process followed first-order kinetics and was adequately described by the equation

$$R_t = R_\infty(1 - \exp(-kt))$$

where R_t and R_∞ represent the activity of ^{42}K in identical volumes of cells at time t and at equilibrium and k is the rate constant of the equilibration process. This equation can be rearranged to give

$$-\ln[1-(R_t/R_\infty)] = kt$$

The figure is a plot of $-\ln[1-(R_t/R)]$ against t during some of the Ca steady states obtained in the experiment of Fig. 4 and is meant to illustrate the method used to calculate k. Since the normal volume–area ratio of the fed red cells remained unaltered during the present experiments, the K permeability (P_K) of the red-cell membrane, can be calculated from k. When $k = 1$ h^{-1}, P_K will be between 1.5×10^{-8} and 2.1×10^{-8} cm s^{-1}. The first numerical value given on each curve corresponds to the Ca^{2+} concentration inside the corresponding cells (in µmol per l cell water).

report here the surprising finding that in the intact, minimally disturbed red cell, Ca activates the K flux with a much lower affinity than that previously found in ATP-depleted red cells or resealed ghosts[1,12,13], suggesting that the higher Ca sensitivity of these systems results from treatment-induced alterations in the reactivity of labile Ca-interacting sites.

To incorporate controlled amounts of Ca inside the cells and to keep a known and constant level of ionised Ca during the measurement of the Ca-induced K flux, we used a divalent cation ionophore, A23187 (ref. 17), and a fast resolution technique for tracer flux measurements, which we developed recently[18] to study cytoplasmic Ca buffering and Ca-pump parameters in intact cells.

Since the normal mammalian red cell is virtually free of Ca^{19}, is almost impermeable to $Ca^{15,16}$ and has a powerful Ca-extrusion pump in its membrane[14], steady states with variable internal Ca concentrations could be obtained only as a balance between active Ca extrusion and ionophore-induced leaks (Fig. 1). The membrane potential was held constant by a 'chemical clamp' method in which the steady-state external concentrations of all permeable ions are in electrochemical equilibrium with the corresponding internal ions at a pre-determined potential, about -10 mV in the present experiments. At each ionophore concentration the level of internal Ca was regulated by varying the Ca concentration in the medium, and the fraction of ionised Ca was estimated from the slope of the curve relating the total internal Ca to the external Ca^{2+} concentration when the Ca pump was saturated (see text of Fig. 1). The influx of K was estimated by measuring the rate of ^{42}K-tracer equilibration at constant internal Ca and with the total external K concentration at electrochemical equilibrium with the internal K (Fig. 2). In every case, single exponentials were obtained with all the internal K participating in the equilibration process. In the absence of Ca, the ionophore had no effect on the Na and K permeability of the membrane. The entry of Ca induced a variable increase in the K flux while at the same time it reduced the Na efflux, presumably by inhibiting the Na pump[1]. Quinine (1 mM) in the medium inhibited the Ca-dependent flux in the present experiments by about 95%, just as it inhibited the Ca-induced increase in K permeability in the absence of the ionophore[3].
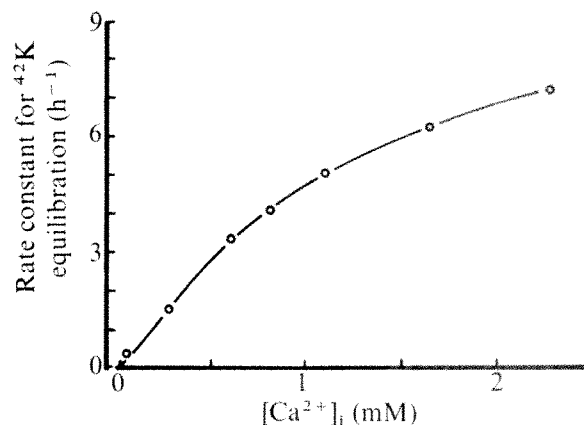
Fig. 3 The rate constant of tracer (^{42}K) equilibration as a function of the internal Ca^{2+} concentration in intact red cells in the presence of a low ionophore concentration. The internal Ca^{2+} concentration and k were obtained as indicated in the legend of Figs 1 and 2 respectively. The proportion of ionophore to cells was about 10 µmol per l of cells.

Figure 3 shows the variation in the rate constant of tracer (^{42}K) equilibration as a function of the internal concentration of Ca^{2+} in intact cells exposed to a relatively low ionophore concentration. The apparent K_m for Ca^{2+} is about 1.2 mM and varied in different experiments with similar ionophore concentrations between 0.33 and 1.5 mM. If the ionophore concentration is increased by ten times or more, the kinetics of Ca activation of the K permeability are completely different (Fig. 4) and approach the pattern previously observed in metabolically depleted resealed ghosts[13]. In ATP-depleted intact cells[10] we always obtained high Ca affinity-type curves, whatever the ionophore concentration, but in such cells the maximum increase in K permeability was only half that of the same cells before depletion. All these different patterns were consistently reproducible.

The variation in Ca sensitivity cannot be attributed to Mg^{2+} since the distribution of this ion at equilibrium must have been the same in all the experiments. To investigate whether the high and low Ca affinity patterns were related to the ATP content

of the cells, we measured the intracellular levels of ATP in the various conditions in which we tested the Ca effect on the K fluxes. The addition of ionophore in the absence of Ca slightly increased the ATP content of the cells in a manner reminiscent of the effect of propranolol[20]. The additional presence of Ca stimulated ATP hydrolysis through the Ca pump and induced a rapid fall in ATP from 1 mM to about 0.55 mM followed by a slower decline, presumably because of the buffering effect of membrane-bound adenylate kinase. These effects were independent of the ionophore concentration, which makes ATP alone an unlikely source of the different kinetic patterns observed in the intact cells. It is also unlikely that the variable Ca sensitivity of the K permeability system results from a non-homogeneous distribution of Ca within the cell since the K fluxes we report here were measured simultaneously in the same cells (see legend to Fig. 1 in ref. 18) where we had previously determined a constant K_m of 1 μM for the Ca activation of the Ca pump at all ionophore concentrations.

We also investigated the effect of ionophore removal on the Ca-dependent K flux. This we could only try in intact red cells loaded with Ca if the Ca pump were inhibited, and for this reason we first depleted the cells of Mg, a necessary internal cofactor for Ca extrusion[21]. Mg extraction and Ca loading were performed in succession by preincubating the cells in the presence of ionophore (10 μmol per 1 cells) first in a medium containing 2 mM EDTA and no Ca or Mg and then in Ca-containing media in the absence of Mg and chelator. The ionophore was then removed by washing the cells eight times in Ca-free ice-cold solutions at low haematocrits (about 1 %). This simple procedure restored the normal Ca impermeability of the intact cell and, with the Ca pump arrested, the cells remained effectively resealed to Ca, losing only about 25 % of their original Ca content during the initial washes. These cells also showed a high Ca affinity pattern similar to that of the intact cells at high ionophore concentrations shown in Fig. 4.

The low Ca affinity pattern shown in Fig. 3 has never been observed before. Since the cells were disturbed least in these conditions we feel that the low affinity response is closer to the behaviour of the normal system. The high affinity pattern would then correspond to an altered reactivity to internal Ca which may or may not occur in the normal cell depending on whether the Ca sensitivity of the K permeability mechanism is or is not subject to control by cellular regulatory mechanisms. Since the maximum increase in K permeability induced by Ca is more than three orders of magnitude larger than the "ground" K permeability of the intact red cell, an increase in Ca_i^{2+} even in the micromolar range[22], may suffice to produce a functionally significant increase in K permeability. A relatively

small effect could thus account for the Ca-induced hyper-polarisation observed in amphiuma red cells[23] and perhaps also in other cells[5,8].

The heterogeneous behaviour of ATP-depleted red cell populations in relation to the Ca-induced K flux[15,23] can now be interpreted as resulting from a variable Ca affinity of the K gate. This revives the old controversy[10] about whether the depleting procedure accumulates[25] or removes[26] metabolites which control the K permeability of red-cell membranes, but now it can be reformulated more precisely in terms of the control of the Ca sensitivity of the K-gating mechanism.

Since many of the conditions in which the effect of internal Ca on the K permeability is obtained resemble those in which 1,2-diacylglycerol accumulates in the red-cell membrane, as recently shown by Allan and Michell[27], we feel tempted to suggest the possibility that 1,2-diacylglycerol, or a similar metabolite, mediates the Ca effects on the K channel. According to Allan and Michell, the rate of formation of this substance depends on the internal Ca^{2+} concentration, whereas the fast conversion to an inactive phosphatide, by the action of a diacylglycerol kinase, would depend on the intracellular level of ATP. If the Ca effect on the K permeability depends on the level of a certain intermediate, the different Ca sensitivities observed in the present experiments may merely reflect the variation in the relative rates of formation and breakdown of this intermediate in the different conditions.

V. L. Lew
H. G. Ferreira

Physiological Laboratory,
University of Cambridge,
Cambridge CB2 3EG, UK

1 Lew, V. L., *J. Physiol., Lond.*, **206**, 35–36P (1970).
2 Gardos, G., *Acta physiol. Hung.*, **10**, 185–193 (1956).
3 Armando-Hardy, M., Ellory, J. C., Ferreira, H. G., Fleminger, S., and Lew, V. L., *J. Physiol., Lond.*, **250**, 32–33P (1975).
4 Ransom, B. R., Barker, J. L., and Nelson, P. G., *Nature*, **256**, 424–425 (1975).
5 Meech, R. W., *Comp. Biochem. Physiol.*, **48**, 387–401 (1974).
6 Clusin, W., Spray, D. C., and Bennett, M. V. L., *Nature*, **256**, 425–427 (1975).
7 Isenberg, G., *Nature*, **253**, 273–274 (1975).
8 Krnjevic, K., *Physiol. Rev.*, **54**, 418–540 (1974).
9 Jenkins, D. M. G., and Lew, V. L., *J. Physiol., Lond.*, **234**, 41–42P (1973).
10 Lew, V. L., *Biochim. biophys. Acta*, **233**, 827–830 (1971).
11 Lew, V. L., *Biochim. biophys. Acta*, **249**, 236–239 (1971).
12 Porzig, H., *Experientia*, **30**, 697–699 (1974).
13 Simons, T., *J. Physiol., Lond.*, **256**, 227–244 (1976).
14 Schatzmann, H. J., *J. Physiol., Lond.*, **235**, 551–596 (1973).
15 Lew, V. L., in *Comparative Biochemistry and Physiology of Transport* (edit. by Bolis, L., Bloch, K., Luria, S. E., and Lynen, F.), 310–316 (North-Holland, Amsterdam, 1974).
16 Ferreira, H. G., and Lew, V. L., *J. Physiol., Lond.*, **252**, 86–87P (1975).
17 Reed, P. W., and Lardy, H. A., *J. biol. Chem.*, **247**, 6970–6977 (1972).
18 Ferreira, H. G., and Lew, V. L., *Nature*, **259**, 47–49 (1976).
19 Schatzmann, H. J., in *Current Topics in Membranes and Transport* (edit. by Bronner, F., and Kleinzeller, A.), 125–168 (Academic, New York, 1975).
20 Manninen, V., *Acta physiol. scand.*, **80**, Suppl. 355, 1–76 (1970).
21 Schatzmann, H. J., and Vincenzi, F. F., *J. Physiol., Lond.*, **201**, 369–395 (1969).
22 Rink, T. J., and Baker, P. F., in *Calcium Transport in Contraction and Secretion* (edit. by Carafoli, E., *et al.*), 235–242 (North-Holland, Amsterdam, 1975).
23 Lassen, U. V., Pape, L., and Vestergaard-Bogind, B., *J. Membr. Biol.*, **18**, 125–136 (1974).
24 Riordan, J. R., and Passow, H., *Biochim. biophys. Acta*, **249**, 601–605 (1971).
25 Passow, H., in *The Red Blood Cell* (edit. by Bishop, C., and Surgenor, D. M.), 113–120 (Academic, New York, 1964).
26 Hoffman, J. F., *Am. J. Med.*, **41**, 666–680 (1966).
27 Allan, D., and Michell, R. H., *Nature*, **258**, 348–349 (1975).

Fig. 4 The rate of ^{42}K equilibration as a function of the internal Ca^{2+} concentration in intact red cells in the presence of a high ionophore concentration. The concentration of ionophore used in this experiment was 1 mg ml^{-1} in the ethanol phase. This is equivalent to a proportion of ionophore to cells of about 100 μmol per 1 of cells.
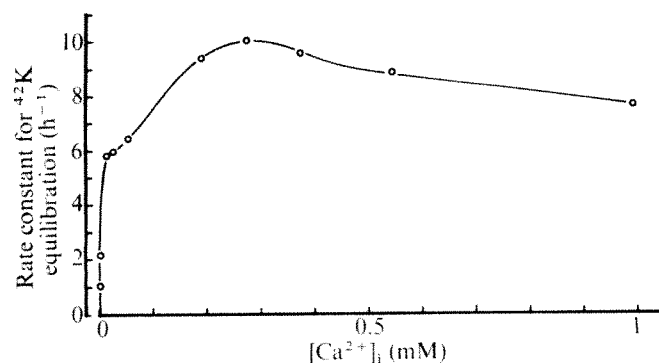
Stereospecificity of interaction of neuroleptic drugs with neurotransmitters and correlation with clinical potency

THE molecular mechanism responsible for the therapeutic activity of a drug is often suggested by the biochemical event which correlates best with clinical potency. Thus, if optical or geometrical isomers of a drug have markedly different clinical potencies, the most relevant biochemical effect should also exhibit isomeric specificity.

While numerous biochemical mechanisms have been proposed to explain the antischizophrenic action of neuroleptics, the

ability of these drugs to antagonise behavioural effects of apomorphine and amphetamine in animals correlates best with clinical potency[1]. Since apomorphine and amphetamine-induced behaviours are mediated by brain catecholamines, especially dopamine, the neuroleptics are thought to act by blocking dopamine receptors in the brain. Clinical potencies of phenothiazine neuroleptics correlate reasonably well with their ability to inhibit a dopamine-sensitive adenylate cyclase which is presumably linked to the dopamine receptor[2,3]. Because pharmacologically potent butyrophenones are relatively weak inhibitors of this cyclase and show only a partial correlation with clinical potency[4], however, it has been suggested that these drugs may only indirectly affect dopamine receptors, perhaps by blocking dopamine release[5]. The chemical similarity of butyrophenones to γ-aminobutyric acid (GABA) suggests that these agents might in some way facilitate GABA transmission in the brain[1]. Since inhibitory GABA neuronal pathways may synapse on dopamine cell bodies, facilitation of GABA effects could reduce the firing rate of dopamine neurones resulting in behaviours like those associated with dopamine receptor blockade[6].

In both clinical and animal pharmacological screening, certain neuroleptics display isomeric specificity. α-Flupenthixol and cis-thiothixene are much more potent clinically and in animals than their geometrical isomers β-flupenthixol and trans-thiothixene respectively[7,8]. Similarly (+)-butaclamol is much more potent than its optical isomer (−)-butaclamol[9]. The relative potencies of butaclamol, flupenthixol and thiothixene isomers upon the dopamine-sensitive adenylate cyclase accord with their isomeric specificity in clinical studies and pharmacological screens[10,11].

Using recently developed binding techniques to study neurotransmitter receptors directly, we have examined critically how reliably stereospecificity and correlations between clinical and biochemical effects can reveal the biochemical basis of neuroleptic drug action. The results suggest that although these criteria are valuable, they can obfuscate rather than clarify drug mechanisms if used incautiously. Of the several biochemical effects of neuroleptics, blockade of dopamine receptors best accounts for therapeutic effects of both butyrophenone and phenothiazine neuroleptics.

Synaptic receptor binding studies were conducted as previously described[12-20]. Briefly, washed crude synaptic membrane or total particulate fractions of various brain regions representing approximately 1 mg of protein were incubated in a buffered solution containing a low concentration of tritiated ligand (1–8 nM) in the presence or absence of various concentrations of the drug under study. After incubation, radioactivity bound to the membranes was separated by filtration or

centrifugation, extracted into an appropriate fluor and counted by liquid scintillation spectrometry. The amounts bound in the presence and absence of the drug were compared. The ligands used in this study and the receptor to which each binds is as follows: ³H-γ-aminobutyric acid (³H-GABA), the GABA receptor[12]; ³H-dopamine (³H-DA and ³H-haloperidol (³H-Halo), the dopamine receptor[13,14]; ³H-naloxone (³H-Nal) ³H-dihydromorphine (³H-DHM), the opiate receptor[15]; ³H-dihydroalprenolol (³H-DHA), the β-adrenergic receptor[16,21]; ³H-5-hydroxytryptamine (³H-5-HT) and ³H-lysergic acid diethylamide (³H-LSD), the 5-HT receptor[17,18]; ³H-quinuclidinyl benzilate (³H-QNB), the muscarinic cholinergic receptor[19]; and ³H-strychnine (³H-Stry), the glycine receptor[20].

Receptor binding for each ligand was assayed in the brain region containing the highest density of receptor sites or in whole brain. The highest density of dopamine receptor sites is in the corpus striatum[13,14]. Binding of ³H-strychnine to glycine is most enriched in spinal cord membranes[20]. Receptor binding for ³H-LSD, ³H-5-HT and ³H-DHA is greatest in the cerebral cortex[16-18,21]. Although there exist regional variations in opiate, GABA and muscarinic receptor binding, the characteristics of binding sites do not vary regionally and substantial activity is present in whole brain or forebrain, which were used for assays of these receptors[12,15,19].

Uptake of radioactive GABA by synaptosomes in nuclei-free sucrose homogenates of whole rat brain was assayed as previously described[22].

All radioactive agents were purchased from New England Nuclear, and all other substances were obtained from the company of origin.

There are no marked stereospecific influences of neuroleptics on receptor binding for ³H-GABA, for the opiates ³H-naloxone and ³H-dihydromorphine, for binding of ³H-strychnine to the glycine receptor, for the binding of ³H-quinuclindinyl benzilate (³H-QNB) to the muscarinic cholinergic receptors or for ³H-dihydroalprenolol (³H-DHA) binding to the β-noradrenaline receptor (Table 1). In contrast, the neuroleptic isomers are extremely potent in competing for ³H-haloperidol binding to the dopamine receptor and display striking stereospecificity. (+)-Butaclamol reduces ³H-haloperidol binding 50% at 1 nM and is more than 1,000 times as potent as (−)-butaclamol in eliciting this effect. α-Flupenthixol inhibits ³H-haloperidol binding 50% at 2 nM and is 50 times as potent as β-flupenthixol, while cis-thiothixene is 100 times more potent than trans-thiothixene in competing for ³H-haloperidol binding. These neuroleptics also demonstrate stereospecificity with respect to ³H-dopamine binding, although they are less potent in competing with this ligand. Thus, while (+)-butaclamol is less than 1% as potent in inhibiting ³H-dopamine as ³H-haloperidol

---

Table 1 Displacement of various receptor ligands by neuroleptic isomers

Compound	³H-DHA[a]	³H-5-HT[b]	³H-LSD[c]	³H-DA[d]	³H-Halo[e]	³H-QNB[f]	³H-Stry[g]	³H-Nal[h]	³H-DHM[i]	³H-GABA[j]
					IC₅₀ (μM)*					
(+)-Butaclamol	N.e.†	1.0	0.05	0.1	0.001	100	N.e.	18	17	N.e.
(−)-Butaclamol	N.e.	8.0	7.0	16.0	1.3	40	N.e.	20	17	N.e.
α-Flupenthixol	N.e.	4.0	0.10	0.2	0.002	0.7	N.e.	55	100	100
β-Flupenthixol	N.e.	60.0	6.0	10.0	0.10	0.9	N.e.	50	100	N.e.
cis-Thiothixene	80	1.0	0.3	0.7	0.003	3.2	50	15	30	50
trans-Thiothixene	80	7.0	1.0	19.0	0.30	2.8	50	15	35	50

*Concentration which inhibits ligand binding 50%.
†No effect at 100 μM.
Four to six concentrations of each drug were evaluated in triplicate to obtain IC₅₀ values. Data are the mean of 2–4 determinations which varied less than 20%.
[a]Rat cerebral cortex membranes, ³H-DHA = 1 nM.
[b]Rat cerebral cortex membranes, ³H-5-HT = 7 nM.
[c]Rat cerebral cortex membranes, ³H-LSD = 3 nM.
[d]Calf striatal membranes, ³H-DA = 5 nM.
[e]Calf striatal membranes, ³H-Halo = 2 nM.
[f]Rat forebrain homogenate, ³H-QNB = 1 nM.
[g]Rat spinal cord–brainstem, synaptosomal membranes, ³H-Stry = 2 nM.
[h]Rat forebrain membranes, ³H-Nal = 1 nM.
[i]Rat forebrain membranes, ³H-DHM = 1 nM.
[j]Rat brain synaptosomal membranes, ³H-GABA = 8 nM.

Table 2 Inhibition of GABA uptake into rat brain synaptosomes by neuroleptic drugs

Compound	IC₅₀ (μM)*	Average clinical daily dose† (μmol kg)
Butyrophenones		
Fluspirilene	3	0.066
Pimozide	10	0.108
Clofuperol	15	0.077
Benperidol	30	0.060
Bromoperidol	30	0.153
Spiroperidol	40	0 058
Penfluridol	60	0.466
Trifluperidol	65	0.096
Haloperidol	75	0.152
Moperone	100	0 802
Droperidol	200	—
Fluanisone	500	3.44
Azaperone	600	—
Pipamperone	1,000	11.10
Phenothiazines and related agents		
Trifluoperazine	17	0 297
Fluphenazine	18	0.168
Triflupromazine	20	4.59
Chlorpromazine	21	12.00
Promazine	30	33.00
α-Flupenthixol	35	0.099
β-Flupenthixol	35	—
cis-Thiothixene	38	0.393
trans-Thiothixene	45	—

*Concentrations which inhibit 50%.
†For clinical potencies, the midpoint values of daily dose ranges[18-21] were meaned and converted to μmol kg⁻¹ assuming a human weight of 70 kg.
Six concentrations of each drug were evaluated in triplicate to obtain IC₅₀ values. Data are the mean of three determinations which varied less than 15%.

binding, it is 160 times more potent than (−)-butaclamol in competing for ^{3}H-dopamine binding. Similarly α-flupenthixol and cis-thiothixene are less than 1% as potent in inhibiting ^{3}H-dopamine as ^{3}H-haloperidol binding yet they still display stereospecificity with respect to their corresponding isomers. Interestingly flupenthixol and thiothixene are just as potent in competing for ^{3}H-QNB as for ^{3}H-dopamine binding, but they display no stereospecificity in inhibiting ^{3}H-QNB binding.

Relative potencies of numerous neuroleptics as inhibitors of ^{3}H-haloperidol binding to the dopamine receptor correlate well with clinical efficacy[14,22,24]. There is, however, a much weaker correlation of neuroleptic clinical and pharmacological potency with ^{3}H-dopamine binding, in spite of the fact that both dopamine and haloperidol presumably interact with the same receptor. The differences between interactions of ^{3}H-haloperidol and ^{3}H-dopamine with the dopamine receptor seem to stem from their labelling two different states of the dopamine receptor. Agonists are much more potent in competing for ^{3}H-dopamine than ^{3}H-haloperidol binding. Conversely, antagonists have substantially higher affinity for haloperidol than dopamine-binding sites. Thus dopamine and haloperidol respectively label discrete "agonist" and "antagonist" states of the dopamine receptor[14]. This two-state model for the dopamine receptor accords with data supporting such a model for the 5-HT, α-noradrenergic and opiate receptors[25]. One would expect pharmacological activities of antagonists to be related better to their affinities for antagonist than agonist states of the receptor, which explains the better correlation of neuroleptic clinical activity with affinity for haloperidol than for dopamine binding[22].

The binding of ^{3}H-LSD and ^{3}H-5-HT is also inhibited stereospecifically and with considerable potency by these neuroleptics. The neuroleptics display about the same potency in inhibiting ^{3}H-LSD as ^{3}H-dopamine binding as well as a similar degree of stereospecificity. Thus (+)-butaclamol is 140 times as potent as (−)-butaclamol in lowering ^{3}H-LSD binding and is about 160 times as potent as the (−)-isomer in lowering ^{3}H-DHA binding. The neuroleptics are somewhat less

potent in lowering the binding of ^{3}H-5-HT, but still display a degree of stereospecificity. The potencies of these neuroleptics in inhibiting ^{3}H-dopamine and ^{3}H-LSD binding are similar to their potencies as inhibitors of the dopamine-sensitive adenylate cyclase[1,2]. In spite of these stereospecific effects, blockade of 5-HT receptors is probably not a major source of neuroleptic clinical efficacy, because relative influences of an extensive series of these drugs on ^{3}H-LSD binding do not correlate with clinical potency[17].

Since the chemical structure of butyrophenones resembles that of the inhibitory neurotransmitter GABA[1], it was also of interest to evaluate the influence of butyrophenones on GABA uptake by synaptosome-enriched brain homogenates. In this type of tissue preparation ^{3}H-GABA is accumulated primarily by nerve terminals[16,17]. The butyrophenones examined are all capable of inhibiting ^{3}H-GABA uptake (Table 2). The most potent, fluspirilene, reduces uptake 50% at 3 μM, similar to its potency in inhibiting ^{3}H-dopamine binding and in inhibiting the dopamine-sensitive adenylate cyclase[1,2]. There is a several hundredfold variation in the relative potencies of the butyrophenones in reducing GABA uptake, and these variations correlate well with the relative clinical potencies of the drugs ($r = 0.86$, $P < 0.001$). If inhibition of GABA uptake were a common mechanism whereby therapeutic effects of all neuroleptics were elicited, then the phenothiazines and related agents should also show a similar correlation between inhibition of GABA uptake and clinical potency. We detect no correlation, however, between the influence of these agents on GABA uptake and their clinical effects. The very potent neuroleptic α-flupenthixol is no more effective in reducing GABA uptake than the weak neuroleptic promazine. Similarly, the potent neuroleptic fluphenazine has about the same potency as the moderately effective neuroleptic chloropromazine.

The most striking finding indicating that the pharmacological activity of neuroleptics related to phenothiazine cannot be attributed to influences on GABA uptake derives from experiments with stereoisomers. No difference is apparent in the potencies of α- and β-flupenthixol in competing for GABA uptake in spite of the fact that essentially all the pharmacological activity resides in the α-isomer with β-flupenthixol being essentially inactive[8]. Similarly, the cis- and trans-isomers of thiothixene have the same affinities for GABA uptake sites, though only cis-thiothixene has clinical and pharmacological activity[7].

It could be argued that the correlation between relative potencies of butyrophenones in inhibiting GABA uptake and their clinical potencies indicates that the therapeutic actions of butyrophenones may be mediated at least in part by inhibition of GABA uptake. Since uptake of GABA into nerve terminals is thought to inactivate synaptically released GABA, such an effect would potentiate the synaptic effects of GABA, which in turn could inhibit the firing of dopamine neurones Similarly inhibition of opiate receptor binding by butyrophenones but not by phenothiazines correlates well with clinical antischizophrenic activity ($r = 0.78$, $P < 0.01$ with no added Na⁺; $r = 0.61$, $P < 0.05$ with 100 mM Na⁺)[23]. There is also some correlation between antipsychotic potencies of neuroleptics and inhibition of dopamine release[6]. Before the identification of ^{3}H-haloperidol binding to dopamine receptor sites[14,22,24] these formulations might have represented plausible mechanisms for clinical effects of butyrophenones if not of phenothiazines. But since the butyrophenones are about 100–1,000 times more potent in competing for ^{3}H-haloperidol binding to dopamine receptor sites than in reducing GABA uptake, opiate binding or dopamine release and since inhibition of haloperidol binding predicts phenothiazine as well as butyrophenone activities, a more parsimonious explanation attributes clinical actions of both butyrophenones and phenothiazines to blockade of dopamine receptors[14,22,24].

The correlation between the influences of these drugs on ^{3}H-GABA uptake and clinical potencies emphasises the difficulties inherent in evaluating correlations between bio-

chemical and clinical effects of drugs. Similarly, one should be circumspect about concluding that stereospecificity of biochemical effects "proves" that one is dealing with the pharmacologically relevant site of a drug's action. Thus, if a given behavioural effect in animals is elicited more by (+)-butaclamol than by (−)-butaclamol, it should not automatically be assumed that the behaviour is dopamine mediated, since the drug effect might also involve 5-HT receptors.

S. J. ENNA
J. P. BENNETT, JR
D. R. BURT
I. CREESE
SOLOMON H. SNYDER

Departments of Pharmacology and Experimental Therapeutics, and Psychiatry and Behavioural Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

1 Janssen, P. A., and Van Bever, W. F., in *Current Developments in Psychopharmacology*, 2 (edit. by Essman, W. E., and Valzelli, L.), 165–184 (Spectrum Publications, New York, 1975).
2 Clement-Cormier, Y. C., Kebabian, J. W., Petzold, G. L., and Greengard, P., *Proc. natn. Acad. Sci. U.S.A.*, 71, 1113–1117 (1974).
3 Miller, R. J., Horn, A. S., and Iversen, L. L., *Molec. Pharmac.*, 10, 759–766 (1974).
4 Iversen, L. L., Rogawski, M. J., and Miller, R. J., *Molec. Pharmac.*, 12, 217–224 (1976).
5 Seeman, P., and Lee, T., *Science*, 188, 1217–1219 (1975).
6 Anden, N.-E., *Archs Pharmac.*, 283, 419–424 (1974).
7 Weismann, A., in *The Phenothiazines and Structurally Related Drugs* (edit. by Forrest, I. S., Carr, C. J., and Usdin, E.), 471–480 (Raven, New York, 1974).
8 Møller-Neilsen, I., Pedersen, V., Nymark, M., Franck, K. F., Boeck, V., Fjalland, B., and Christensen, A. V., *Acta pharmac. tox.*, 33, 353–362 (1973).
9 Bruderlein, F. T., Humber, L. G., and Voith, K., *J. med. Chem.*, 18, 185–188 (1975).
10 Lippman, W., Pugsby, T., and Marker, J., *Life Sci.*, 16, 213–244 (1975).
11 Iversen, L. L., *Science*, 188, 1082–1089 (1975).
12 Enna, S. J., and Snyder, S. H., *Brain Res.*, 100, 81–97 (1975).
13 Burt, D. R., Enna, S. J., Creese, I., and Snyder, S. H., *Proc. natn. Acad. Sci. U.S.A.*, 72, 4655–4659 (1975).
14 Creese, I., Burt, D. R., and Snyder, S. H., *Life Sci.*, 17, 993–1002 (1975).
15 Pasternak, G. W., Snowman, A. M., and Snyder, S. H., *Molec. Pharmac.*, 11, 735–744 (1975).
16 Bylund, D. B., and Snyder, S. H., *Molec. Pharmac.*, 12, 568–580 (1976).
17 Bennett, J. P., Jr and Snyder, S. H., *Brain Res.*, 94, 523–544 (1975).
18 Bennett, J. P., Jr, and Snyder, S. H., *Molec. Pharmac.*, 12, 373–389 (1976).
19 Yamamura, H. I., and Snyder, S. H., *Proc. natn. Acad. Sci. U.S.A.*, 71, 1725–1729 (1974).
20 Young, A. B., and Snyder, S. H., *Proc. natn. Acad. Sci. U.S.A.*, 70, 2832–2836 (1973).
21 Alexander, R. W., Davies, J. N., and Lefkowitz, R. J., *Nature*, 258, 437–440 (1975).
22 Bennett, J. P., Logan, W. J., and Snyder, S. H., *Science*, 178, 997–999 (1972).
23 Creese, I., Burt, D. R., and Snyder, S. H., *Science*, 192, 481–483 (1976).
24 Seeman, P., Chau-Wong, M., Tedesco, J., and Wong, K., *Proc. natn. Acad. Sci. U.S.A.*, 72, 4376–4380 (1975).
25 Snyder, S. H., and Bennett, J. P., Jr, *A. Rev. Physiol.*, 38, 153–175 (1976).
26 Iversen, L. L., and Bloom, F. E., *Brain Res.*, 41, 131–134 (1972).
27 Hokfelt, T., and Ljungdahl, A., *Expl Brain Res.*, 14, 354–363 (1972).
28 Usdin, E., and Efron, D. H., *Psychotropic Drugs and Related Compounds* (US Department of Health Education and Welfare, Washington, DC, 1972).
29 Mielke, D. H., Gallant, D. M., Oelsner, T., Kessler, C. M., Tomlinson, W. K., and Cohen, G. H., *Dis. Nerv. Syst.*, 36, 7–8 (1975).
30 Ayd, F. J., *Int. Drug Therapy Newsletter*, 6, 17–20 and 25–30 (1971); 7, 13–16 (1972); 10, 25–28 (1975).
31 Byck, R., in *The Pharmacological Basis of Therapeutics* (edit. by Goodman, L. S., and Gilman, A.), 152–200 (MacMillan, New York, 1975).
32 Creese, I., Feinberg, A., and Snyder, S. H., *Eur. J. Pharmac.*, 36, 231–235 (1976).

Transmitter metabolism in substantia nigra after inhibition of dopaminergic neurones by butyrolactone

DOPAMINERGIC nigrostriatal neurones have recently been found to contain dopamine (DA) also in their dendrites[1]. Certain observations suggest a release of transmitter from these dendrites: depolarisation by potassium ions induces a release of ^{3}H-DA from incubated tissue slices of rat substantia nigra[2], and stimulation of the median forebrain bundle increases 3,4-dihydroxyphenylacetic acid, a major metabolic of DA[3], in the mesencephalic region containing dopaminergic cell bodies and dendrites[4], as well as in caudate–putamen, the terminal area. These similarities

between terminal and somato-dendritic areas point to a similar regulation of DA metabolism and an active role of DA in both parts of the neurone. We have studied this question in a situation where the dopaminergic neurones are inhibited. γ-Butyrolactone (GBL) applied systemically, strongly reduces unit activity of the dopaminergic neurones in zona compacta of substantia nigra[5]. Under these conditions, the changes observed in the somato–dendritic complex differed markedly from those reported for the terminal area[6-12].

In male rats (200–300 g) of a random-bred SIV (Sprague–Dawley–Ivanovas) strain, the somato–dendritic complex was analysed as a whole with biochemical techniques, while DA nerve cell bodies were studied by histochemical microfluorimetry[13]. For biochemical analyses the brains were quickly removed after decapitation, placed on dry ice and covered with carbon dioxide snow. Standardised tissue cylinders (about 1 mg) from substantia nigra were obtained as described in Fig. 1[14]. These cylinders contained the medial two thirds of zona compacta and small amounts of tissue from zona reticulata and lemniscus medialis. Their localisation corresponds to the area of highest dendritic density[1] and to the highest DA concentration in substantia nigra (F. Hefti and W. Lichtensteiger, unpublished). DA concentration was determined in individual tissue cylinders with an enzymatic–isotopic method[15]. The rate of DA synthesis was assessed by measuring accumulation of dopa after inhibition of dopa-decarboxylase with 3-hydroxybenzylhydrazine (NSD 1015, Sandev) with a newly developed enzymatic–isotopic assay[16]. In these conditions dopa accumulated linearly in substantia nigra cylinders up to 60 min (ref. 16).

Both inhibition of dopaminergic cells by GBL[6,8,9] and interruption of impulse flow by axotomy[9,10,12] have been found to lead to an increase in the concentration of dopamine in caudate–putamen. After a single injection of 750 mg kg⁻¹ of GBL. DA concentration was greatly increased between 15 and 150 min (ref. 8). We could confirm
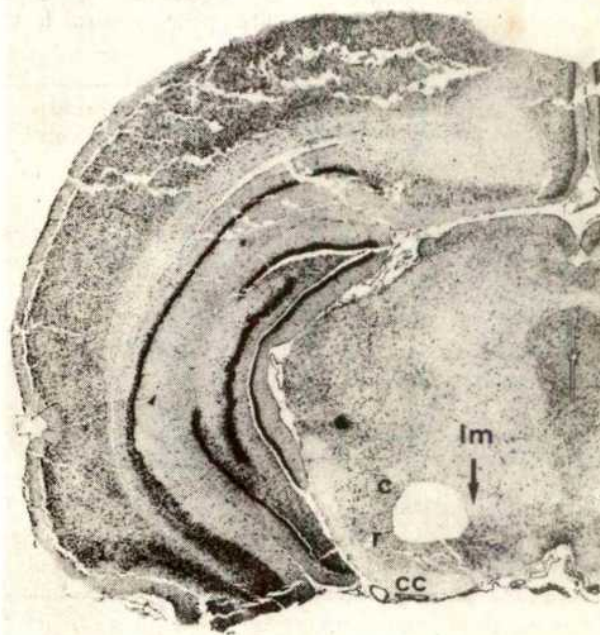
Fig. 1 Localisation of standardised tissue cylinders in substantia nigra. From frozen brain, a 1-mm thick frontal slice was sectioned extending anteroposteriorly from level A 1400 to level A 2400 according to the atlas of Koenig and Klippel[20]. A cylinder of 1-mm diameter was punched out on each side. The figure shows the histological control of such a slice fixed after punching (Toluidine blue staining). c, Zona compacta; r, zona reticulata of substantia nigra; lm, lemniscus medialis; cc, crus cerebri.
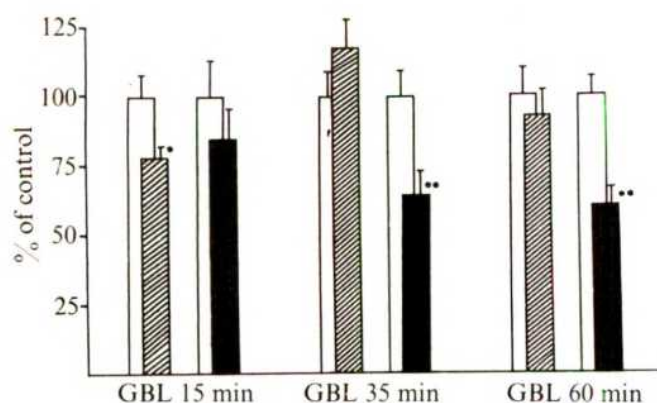
Fig. 2 Effect of γ-butyrolactone (GBL) on DA concentration (hatched bars) and dopa accumulated in substantia nigra 30 min after injection of the dopa-decarboxylase inhibitor NSD 1015 (100 mg kg^{-1} intraperitoneally; solid bars). Open bars represent saline injected controls. (Mean $\pm$s.e.m.; $n = 9$–10; *$P<0.025$; **$P < 0.005$). Concentration of DA and dopa were based on protein content of tissue cylinders[21]. Mean $\pm$s.e.m. of DA control levels was 15.0 ± 1.1 ng per mg protein, mean of dopa accumulated in 30 min was 8.3 ± 0.8 ng per mg protein in control animals.

this effect in rats (at 60 min) and observed a comparable change in mice (at 10 and 90 min; unpublished). In contrast to this the DA concentration in substantia nigra was not raised after 35 and 60 min; there was even a reduction after 15 min (Fig. 2). The quantitative histochemical investigation of DA perikarya in zona compacta of substantia nigra revealed no changes in fluorescence intensity in male rats 60 min after GBL, while intensity tended to decrease in male mice at 10 min and 90 min after drug administration (Table 1). This is again at variance with the response of DA terminals, where GBL elicits a marked rise in fluorescence intensity[7,17].

Also with regard to DA synthesis, our observations on the somato–dendritic area (Fig. 2) contrast with the findings in the terminal region[10,11]. In substantia nigra, dopa accumulation remained about 40% below control level at

35 and 60 min, whereas the amount of dopa accumulated in caudate-putamen has been shown to be more than doubled 35 min after GBL and to be again at control levels after 90 min[10].

Since the decarboxylase inhibitor NSD1015 was reported to slightly inhibit monoamine oxidase[18], it could be argued that the absence of an increase in DA synthesis rate in substantia nigra might be related to effects on intraneuronal feedback mechanisms on tyrosine hydroxylase. The effects of GBL on dopa accumulation in caudate-putamen were studied with another inhibitor of dopa-decarboxylase, N^1-(DL-seryl)-N^2-(2,3,4-trihydroxybenzyl)-hydrazine (Ro 4-4602) (ref. 10). In caudate-putamen, however, an increased accumulation of dopa was found after inhibition of impulse flow by axotomy also with NSD1015 (ref. 11). Moreover, Walters and Roth[19] have shown that prior application of an inhibitor of monoamine oxidase does not prevent the increase in dopa accumulation in caudate-putamen after Ro 4-4602 and GBL, only the magnitude of the effect was slightly reduced. Inhibition of monoamine oxidase for 30 min did not deruce dopa accumulation after Ro 4-4602 alone. Therefore, it seems that the discrepancies between substantia nigra and caudate-putamen are not related to the type of decarboxylase inhibitor used, but rather reflect a true difference between the two regions.

Our findings indicate that somatodendritic and terminal areas of dopaminergic neurones differ with regard to the regulation of monoamine metabolism at least in some functional states. The relative importance of biochemical changes in dendrites and perikarya is as yet difficult to assess. A rise in dendritic DA after GBL may have been obscured by a fall in DA concentration of perikarya, or else, it may have been too small to be detected against the background of perikaryal DA. Our microfluorimetric data speak against the first possibility. Moreover, the contribution of dendrites to nigral DA seems to be considerable, as DA concentration in zona reticulata, which contains mainly dendrites, is relatively high (4.2 ± 0.7 ng per mg protein; mean$\pm$s.e.m.; $n=10$) as compared to zona compacta containing a mixture of perikarya and dendrites (14.9 ± 1.5 ng per mg protein; $n=10$). It follows from these considerations that the dopaminergic dendrites themselves probably responded to GBL in a way different from the reaction of terminals.

F. Hefti
R. Lienhart
W. Lichtensteiger

Department of Pharmacology,
University of Zürich,
Gloriastr. 32A,
CH-8006 Zürich, Switzerland

Table 1 Fluorescence intensity of dopamine-containing nerve cells in zona compacta of substantia nigra*

Treatment	Mean relative fluorescence intensity†	Mean $\pm$ s.d. of logarithmically transformed intensity distribution (natural log)	Cell count (No. of animals)
Male rats			
Saline, 60 min	60.58	4.044 ± 0.3602	779 (5)
GBL 750 mg kg^{-1}, 60 min	62.44	4.063 ± 0.3842	762 (5)
Male mice			
Saline, 10 min	51.32	3.861 ± 0.4150	749 (5)
Saline, 40 min	52.47	3.866 ± 0.4548	803 (5)
GBL 375 mg kg^{-1}, 10 min	48.42	$3.790\pm0.4515\ddagger$	782 (5)
GBL 375 mg kg^{-1}, 90 min	48.11	$3.755\pm0.5382\ddagger$	781 (5)

*The nigral DA neurone group was investigated in 8 frontal 7-μm sections (distance between sections = 35 μm). Twenty neighbouring cells were measured in each section, starting at the dorsolateral corner of zona compacta (for further details, compare refs 13 and 22). The experiments on mice were performed in the course of a study on morphine action; controls are taken from that study[22].
†Mean relative intensity in per cent of noradrenaline standard. For statistical analysis, the values of individual cells were logarithmically transformed. Mean values were compared according to Scheffé's method of linear contrasts[13].
‡Different from both saline controls for $P < 0.05$.

[1] Biörklund, A., and Lindvall, O., *Brain Res.*, **83**, 531–537 (1975).
[2] Geffen, L. B., Jessell, T. M., Cuello, A. C., and Iversen, L. L., *Nature*, **260**, 258–260 (1976).
[3] Roffler-Tarlov, S., Sharman, D. F., and Tegerdine, P., *Br. J. Pharmac.*, **42**, 343–351 (1971).
Wilk, S., Watson, E., and Travis, B., *Eur. J. Pharmac.*, **30**, 238–243 (1975).
[4] Korf, J., Zieleman, M., and Westerink, B. H. C., *Nature*, **260**, 257–258 (1976).
[5] Roth, R. H., Walters, J. R., and Aghajanian, G. K., *Frontiers in Catecholamine Research* (edit. by Usdin, E., and Snyder, S.), 567–574 (Pergamon, New York, 1975).
[6] Gessa, G. L., *et al.*, *Life Sci.*, **5**, 1921–1930 (1966).
[7] Aghajanian, G. K., and Roth, R. H., *J. Pharmac. exp. Ther.*, **175**, 131–138 (1970).
[8] Walters, J. R., and Roth, R. H., *Biochem. Pharmac.*, **21**, 2111–2121 (1972).
[9] Walters, J. R., Roth, R. H., and Aghajanian, G. K., *J. Pharmac. exp. Ther.*, **186**, 1630–1639 (1973).
[10] Walters, J. R., and Roth, R. H., *J. Pharmac. exp. Ther.*, **191**, 82–91 (1974).
[11] Kehr, W., Carlsson, A., Linqvist, M., Magnusson, T., and Atack, C., *J. Pharm. Pharmac.*, **24**, 744–747 (1972).
[12] Stock, G., Magnusson, T., and Andén, N. E., *Naunyn Schmeideberg's Arch. Pharmac.*, **278**, 347–361 (1973).
[13] Lichtensteiger, W., *Prog. Histochem. Cytochem.*, **1**, 185–276 (1970); Lichtensteiger, W., Lienhart, R., Hefti, F., and Felix, D., *Brain Res.* (in the press).
[14] Schlumpf, M., Waser, P. G., Lichtensteiger, W., and Langemann, H., *Biochem. Pharmac.*, **23**, 2447–2449 (1974).
[15] Palkovits, M., Brownstein, M., Saavedra, J. M., and Axelrod, J., *Brain Res.*, **77**, 137–149 (1974).

16 Hefu, F., and Lichtensteiger, W., *J Neurochem*, 27, 647–649 (1976).
17 Emarson, P., Hallman, H, and Jonsson, G, *Med Biol.*, 53, 15–24 (1975)
18 Carlsson, A, Kehr, W, Lindqvist, M., Magnusson, T., and Atack, C V., *Pharmac. Rev*, 24, 371–384 (1972)
19 Walters, J. R, and Roth, R. H, *Biochem Pharmac.*, 25, 649–654 (1976)
20 Koenig, J F. R., and Klippel, R. A, *The Rat Brain. A Stereotaxic Atlas* (Williams and Wilkins, Baltimore, 1963)
21 Lowry, O H., Rosebrough, N. J, Farr, A L, and Randall, R. J, *J biol Chem.*, 193, 265–275 (1951)
22 Lienhart, R., Lichtensteiger, W, and Langemann, H, *Naunyn-Schmiedeberg's Arch. Pharmac*, 286, 353–369 (1975).

Technique for studying synaptic connections of single motoneurones in man

DURING normal voluntary muscle contractions, the pattern of discharge of a single motoneurone is determined by the summed effects of an enormous variety of different types of synaptic input. In principle, the effect of any one of these inputs could be determined by constructing a histogram of the time of occurrence of motoneurone spikes after repeated presentation of a suitably controlled stimulus[1,2]. This procedure extracts from the naturally occurring spike train only those changes in firing time locked to the stimulus. The synaptic effect of a given input is thus revealed in terms of the specific contribution it makes to the total firing pattern of the cell. Such studies are

Fig. 1 *a*, Post-stimulus time histogram of the occurrence of the action potential of a single motor unit in the first dorsal interosseous muscle after a sharp tap to the tip of the nail of the index finger. Stimuli delivered at $3 s^{-1}$ while the subject maintained an isometric contraction such that the unit fired steadily at about 8 p.p.s. 4,096 sweeps. Bin width 1 28 ms. Stimulus given at time zero. *b*, Twitch response of the same motor unit recorded by averaging the force of abduction of the index finger exerted at the second proximal interphalangeal joint triggering the averager from the motor unit action potential while the subject maintained a contraction such that the unit fired at < 3 p p s For details of method see refs 3 and 4. 64 sweeps. Threshold of recruitment 160 g.

most conveniently carried out in man where the cooperation of the subject is invaluable. This report describes an example of this approach in the study of the synaptic effects of a simple cutaneous input. In more general terms, the experimental situation described enables, for the first time, the synaptic connections of single motoneurones to be examined in man, studies previously only attempted in anaesthetised or decerebrate animals.

Figure 1*a* shows the changes in probability of firing of a single motor unit in a hand muscle after a sharp mechanical tap to the tip of a fingernail. At 39 ms after the tap there is an increase in probability of occurrence of the motor unit action potential which reaches a maximum 7 ms later. This period of excitation is immediately followed by a period of inhibition during which the probability of motor unit firing reaches a minimum at 58 ms. This itself is terminated by another more pronounced and prolonged period of excitation reaching a

maximum at 76 ms. This is followed by some rather slow fluctuations superimposed on a prolonged period of reduced probability which eventually returns to control about 300 ms after the initial stimulus[5]. The timing of these responses suggests that the earliest excitation and inhibition involve spinal pathways and that the later more pronounced excitation involves supraspinal pathways[6,7].

The possibility that these responses might be due to vibration transmitted to muscle receptors has been tested by anaesthetising the index finger. The responses can be abolished either by a 2% Lignocaine or by an ischaemic block of the digital nerves. Moreover, weak electrical stimulation of the digital nerves at two times threshold for perception gives essentially the same result as tapping the nail.

The scale of the response to such a modest and limited stimulus, illustrates the most powerful influence that a cutaneous afferent volley can exert on the discharge of a motoneurone. During the initial inhibitory phase, the unit is almost completely silenced. Immediately after this, the mean firing rate is more than doubled. It is clear that the results of experiments in which the response of the motor system is studied after sudden external disturbance must be interpreted with caution if any sudden displacements of the skin are involved[8].

The results of the present type of experiment gain additional significance because the reflex connections of a motoneurone can be related to its recruitment threshold during voluntary muscle contraction and to the mechanical properties of the muscle fibres it innervates (Fig. 1*b*). It will now be possible, in man, to test the applicability of the unexpected subtleties revealed by intracellular recording in the synaptic connections of cat hind limb motor units with different mechanical properties[9,10].

J. A. STEPHENS
T. P. USHERWOOD
R. GARNETT

Sherrington School of Physiology,
St Thomas's Hospital Medical School,
London SE1 7EH, UK

[1] Moore, G. P., Perkel, D. H., and Segundo, J. P., *A. Rev. Physiol.*, **28**, 493–522 (1966).
[2] Perkel, D. H., Gerstein, G. L., and Moore, G. P., *Biophys. J.*, **7**, 391–418 (1967).
[3] Milner-Brown, H. S., Stein, R. B., and Yemm, R., *J. Physiol., Lond.*, **228**, 285–306 (1973).
[4] Stephens, J. A., and Usherwood, T. P., *J. Physiol., Lond.*, **250**, 37–38P (1975).
[5] Stephens, J. A., and Usherwood, T. P., *J. Physiol., Lond.*, **258**, 49P–51P (1976).
[6] Caccia, M. R., McComas, A. J., Upton, A. R. M., and Blogg, T., *J. Neur. Neurosurg. Psychiat.*, **30**, 960–977 (1973).
[7] Godaux, E., and Desmedt, J. E., *Brain Res.*, **85**, 447–458 (1975).
[8] Marsden, C. D., Merton, P. A., and Morton, H. B., *Nature*, **238**, 140–143 (1972).
[9] Burke, R. E., in *New Developments in Electromyography and Clinical Neurophysiology*, 3 (edit. by John E. Desmedt), 69–94 (Karger, Berlin, 1973).
[10] Burke, R. E., Rymer, W. Z., and Walsh, J. V., in *Control of Posture and Locomotion* (edit. by Stein, R. B., Pearson, K. G., Smith, R. S., and Redford, J. B.), 29–44 (Plenum, London, 1973).

Voltage-dependent action of tetrodotoxin in mammalian cardiac muscle

TETRODOTOXIN (TTX) in many excitable membranes selectively blocks the transient change in sodium conductance which underlies the rising phase of the action potential. Electrophysiological and binding studies from nerve and skeletal muscle[1–5] indicate that a one-to-one binding reaction between toxin and Na channels with a dissociation constant of about 5×10^{-9} M can adequately describe the kinetic and steady-state actions of TTX on the Na conductance. Cardiac muscle is unusual in that it shows a marked insensitivity to TTX. A substantial reduction in the rate of rise of action potentials in ventricular trabeculae and Purkinje fibres does not occur until a TTX concentration of about 10^{-6} M is reached[6–8]. We report here the results of a systematic study of the effects of TTX on mammalian cardiac muscle. The experiments were designed to determine whether the reported difference in sensitivity to TTX is an indication of basic differences in the TTX receptor of cardiac muscle as compared to nerve and skeletal muscle. Unlike any TTX effects reported in other tissues[6], our results show a voltage dependence of TTX block. This suggests possible differences in the structure of the TTX binding site or of the Na channel itself.

Intracellular membrane action potentials were recorded from guinea pig papillary muscles by conventional microelectrode techniques, and the maximal upstroke velocities ($\dot{V}_{max}$) of the action potentials were differentiated electronically. The papillary muscles were mounted in a single sucrose gap arrangement as described previously[9]. The sucrose gap permitted uniform displacement of the membrane potential in the short (< 0.8 mm) muscle end where action potentials were recorded. The muscles were stimulated with brief current pulses (5 ms; 0.3 Hz), and

occasionally long lasting depolarising or hyperpolarising constant currents were applied. Our experiments showed that $\dot{V}_{max}$ was linearly related to the extracellular Na concentration (correlation coefficient, $r = 0.99$), indicating that the corresponding membrane current is carried by Na ions as in other excitable tissues. The blocking action of TTX on $\dot{V}_{max}$ reached a steady state in about 5 min, and no further effect of the toxin was seen after continuous exposure for up to 2 h. For concentration–response curves, $\dot{V}_{max}$ measurements were typically made after 15 min equilibration in solutions containing toxin and can be considered as steady state. Continuous microelectrode impalements in a single cell were routinely maintained throughout an experiment allowing a direct comparison of the effects of several toxin concentrations on a given cell at two membrane potentials. Membrane resting potential, V_r, was either changed by increasing the potassium concentration in normal Tyrode's solution (5.4 K⁺-Tyrode's) to 10.8 mM (10.8 K⁺-Tyrode's) or by passing constant currents across the sucrose gap.

The concentration dependence of TTX inhibition of $\dot{V}_{max}$ is shown in Fig. 1. The curves were fitted by assuming a first-order reaction between toxin and receptor. The half-maximal inhibition of $\dot{V}_{max}$ in 5.4 K⁺-Tyrode's occurs at 1.4×10^{-5} M TTX. Doubling the potassium concentration of the Tyrode's solution profoundly affects the sensitivity of the muscle to TTX. Only 1.2×10^{-6} M TTX is required for half-maximal inhibition of $\dot{V}_{max}$ in 10.8 K⁺-Tyrode's. The high potassium solution depolarised the fibres from -89 ± 3 mV to -71 ± 3 mV ($\pm$s.e.). The increase in apparent affinity to TTX is a result of this depolarisation and not of the increased K⁺ concentration *per se*, since a depolarisation of similar magnitude produced by passing a constant current through the fibre had a similar effect on the TTX sensitivity.

$\dot{V}_{max}$ is an indirect and nonlinear measure of the maximal sodium conductance g_{Na} (refs 3 and 10) which is actually the parameter directly affected by TTX binding[1,3,6]. We thought it possible that the depolarisation induced decrease in the control value of $\dot{V}_{max}$ (Figs 1 and 2) might change the relationship between $\dot{V}_{max}$ and g_{Na} (ref 10), which in turn could alter the apparent sensitivity to TTX. Therefore, we carried out experiments with 5.4 K⁺-Tyrode's solution in which 50% of the Na⁺ had been replaced with choline⁺. In these conditions $\dot{V}_{max}$ is reduced by half—that is, even more than in 10.8 K⁺-Tyrode's. There was a small, non-significant ($P > 0.1$) difference in the concentration–response curves in normal and low Na⁺-Tyrode's which could not explain the depolarisation-induced
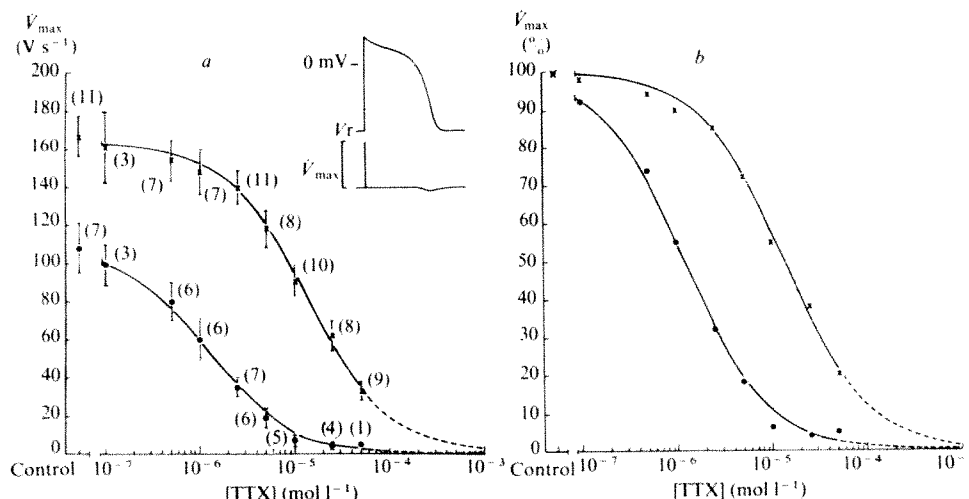
Fig. 1 *a*, Concentration–response curves showing the effects of TTX (log scale) on maximum upstroke velocity ($\dot{V}_{max}$; mean $\pm$ s.e.) of the action potential in 5.4 ($\times$) and 10.8 ($\bullet$) mM K⁺-Tyrode's. The inset is a schematic drawing of resting potential (V_r) and action potential (upper trace); the latter was differentiated electronically (lower trace) and the maximal upstroke velocity ($\dot{V}_{max}$) was measured. *b*, Normalised mean values of *a*. Solid lines are concentration–response curves calculated from the equation

$$\dot{V}_{max\ TTX} = \dot{V}_{max\ control}\,(1 - [1/(1 + \{K_m/[TTX]\})])$$

$\dot{V}_{max\ control}$ and K_m were 167 V s⁻¹ and 1.4×10^{-5} mol l⁻¹ in 5.4 K⁺-Tyrode's, and 109 V s⁻¹ and 1.2×10^{-6} mol l⁻¹ in 10.8 K⁺-Tyrode's. Numbers of determinations per mean value given in parentheses in *a*.

increase in sensitivity to TTX. We do not yet know whether the measurements of V_{max} reflect directly the true binding behaviour of TTX to the Na channels. Studies on skeletal muscle, however, show that measurements of V_{max} give values for the apparent dissociation constant of TTX which agree within factors of 2–6 with those obtained from binding studies[4,5,11].

The concentration–response curves in Fig. 2 were taken from a single papillary muscle which was treated with various TTX concentrations in 10.8 and 5.4 K$^+$-Tyrode's. Initially the preparation was superfused with 10.8 K$^+$-Tyrode's and the concentration of TTX was increased until V_{max} was completely
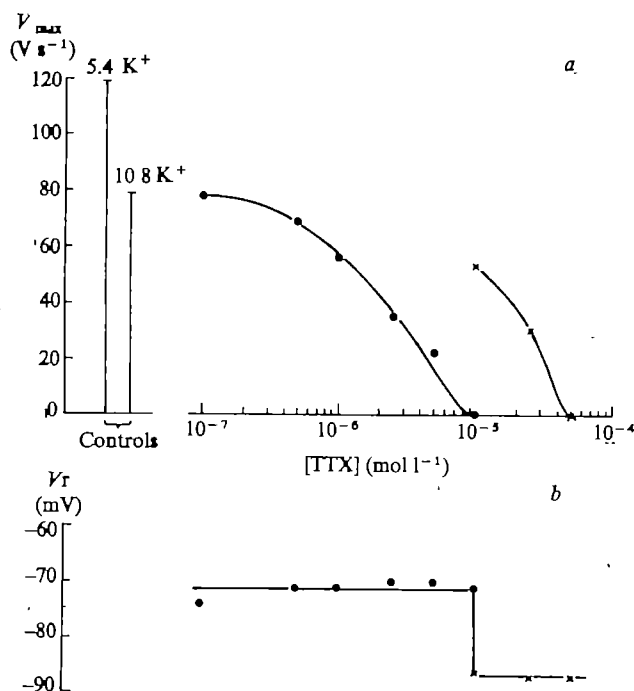
Fig. 2 Concentration–response curves of TTX on V_{max} in 10.8 (●) and 5.4 (×) mM K$^+$-Tyrode's. Data were collected from one papillary muscle during a single microelectrode impalement. *a*, Bars on the left give control values of V_{max} before application of TTX. Initially, the preparation was bathed in 10.8 K$^+$-Tyrode's and the TTX concentration (log scale) increased until V_{max} fell to zero; 5.4 K$^+$-Tyrode's containing 10^{-4} mol l^{-1} TTX was then applied and V_{max} returned to 45% of its normal value and fell again to zero upon further increase in the TTX concentration. *b*, Changes in the resting potential (V_r) caused by the variation in potassium concentration.

abolished, even after increasing the stimulus strength. When the solution was changed to 5.4 K$^+$-Tyrode's containing the same amount of toxin that had caused a complete block of V_{max} in 10.8 K$^+$-Tyrode's, a rapid recovery of V_{max} to about 50% of its control value occurred, while V_r hyperpolarised by 16 mV. When the TTX concentration was increased further, the fibre again showed the lower sensitivity to toxin expected at this V_r. One possible explanation of this effect is that the Na channels inactivated during depolarisation are protected from TTX binding, but become sensitive to TTX again on hyperpolarisation when inactivation is removed.

We have tested this possibility by incubating non-stimulated muscles for 10 min with TTX (10^{-4} M) at −85 mV where all Na channels are available and at −50 mV (20 K$^+$-Tyrode's) where they are fully inactivated[9,13]. When the muscles were rapidly hyperpolarised by passing a constant current and stimulated from a V_r of about −90 mV, the degree of inhibition of V_{max} was slightly larger following the period of inactivation. This suggests that the depolarisation induced inactivation of the Na channels did not protect them from the TTX action. Therefore, we explain the result in Fig. 2 in terms of a release

of TTX during hyperpolarisation to −87 mV from Na channels which were blocked by the toxin at −71 mV. In other words, the hyperpolarisation partially removed the TTX block attained at the lower V_r. This argues strongly for voltage dependence of TTX binding.

Our results indicate at least two major differences between the action of TTX in cardiac muscle and in nerve or skeletal muscle. First, V_{max} of action potentials of cardiac muscle is three to four orders of magnitude less sensitive to the toxin. Second, the sensitivity of cardiac fibres to TTX is voltage dependent, a depolarisation of approximately 18 mV from the normal V_r being sufficient to change the apparent affinity by a factor of 10. The voltage-dependent effects of TTX on cardiac muscle indicate to us a basic difference in the chemistry or configuration of the TTX receptor, or of the Na channel[13]. Possible differences in the structure of the Na channel in cardiac muscle as compared with other excitable tissues are relevant to the mechanism of action of antiarrythmic drugs, which primarily affect the Na conductance system.

This work was supported by the Swiss NSF.

M. Baer
P. M. Best
H. Reuter

Department of Pharmacology,
University of Bern,
Friedbuehlstrasse 49,
3010 Bern, Switzerland

1 Hille, B , *J. gen. Physiol* , 51, 199–219 (1968).
2 Colquhoun , D . Henderson , R , and Ritchie, J. M., *J Physiol.*, *Lond.*, 227, 95–126 (1972).
3 Schwarz, J R., Ulbricht, W , and Wagner, H. H., *J Physiol.*, *Lond.* 233, 167–194 (1973)
4 Almers, W , and Levinson, S. R., *J. Physiol.*, *Lond* , 247, 483–509 (1975).
5 Jaimovich, E , Venosa, R. A., Shrager, P , and Horowicz, P., *J gen. Physiol.*, 67, 399–416 (1976).
6 Narahashi, T , *Physiol. Rev* , 54, 813–889 (1974).
7 Scholz, H., *Naunyn-Schmiedeberg's Arch. Pharmac.*, 265, 187–204 (1969)
8 Dudel, J., Peper, K , Rudel, R., and Trautwein, W , *Pflügers Arch.*, 295, 213–226 (1967).
9 Gettes, L. S., and Reuter, H., *J Physiol.*, *Lond.*, 240, 703–724 (1974)
10 Hunter, P. J., Mc Naughton, P A , and Noble, D , *Prog. Biophys. molec. Biol.*, 30, 99–144 (1975).
11 Colquhoun, D , Rang, H. P., and Ritchie, J. M., *J. Physiol.*, *Lond* , 240, 199–226 (1974).
12 Booker, G. W , Jr, and Reuter, H , *J. Physiol* , *Lond* , 207 165–190 (1970)
13 Hille, B , *Biophys J* , 15, 615–619 (1975)

Localisation of gonadotropin-releasing and thyrotropin-releasing hormones in human brain by radioimmunoassay

Gonadotropin-releasing hormone (GnRH) and thyrotropin-releasing hormone (TRH) are two hypothalamic oligopeptides which regulate the secretion of gonadotropins and of thyroid-stimulating hormone (TSH), respectively, from the anterior pituitary[1,2]. It is generally accepted that these neurohormones are produced by nerve cells, transmitted along axons and stored in nerve endings; when released, they are transferred by way of the hypothalamo–hypophyseal portal system to the anterior pituitary[3]. Human brains of adults have been shown to contain GnRH and TRH activity[4] but quantitative measurements of these hormones are available for brains of aborted foetuses only[5]. This report deals with the distribution of GnRH and TRH in the adult human brain, their possible sites of formation, routes of migration and sites of storage in the pituitary stalk.

The hypothalami and other brain regions were removed during routine autopsies carried out at the Institute of Forensic Medicine, Tel Aviv. The tissues were obtained from three 22–40-yr-old females and twelve 41–95-yr-old males, most of them killed in road accidents. The bodies were kept for 6–48 h at 4 °C until autopsy. The pituitary stalk and gland were dissected
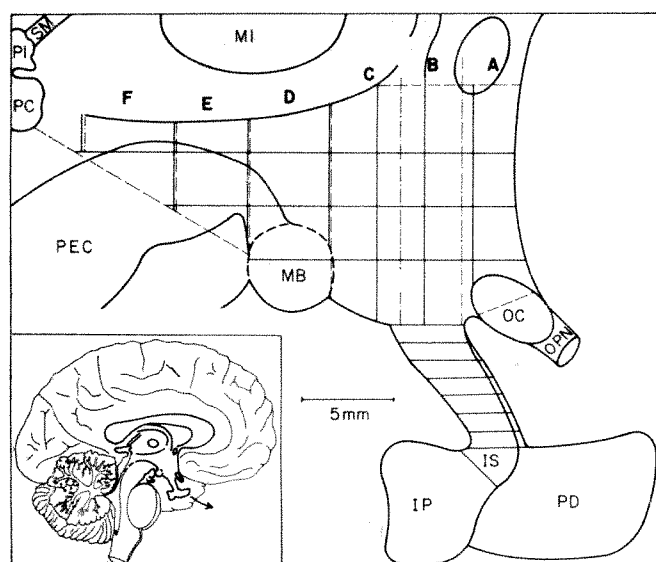
Fig. 1 Outline of the dissection of the human hypothalamus and pituitary stalk. Hypothalamus was dissected in the coronal planes to six strips (broken lines, A–F) or into 25 blocks (solid lines). The pituitary stalk was sectioned into eight portions (including the infundibular stem). Insert: mid-sagittal section of the human brain. AC, Anterior commissure; CF, column of fornix; DM, dorsomedial nucleus; HS, hypothalamic sulcus; IN, infundibular nucleus; INF, infundibulum of neurohypophysis; IP, infundibular process of neurohypophysis; IS, infundibular stem of neurohypophysis; MB, mammillary body; MI, massa intermedia; OC, optic chiasma; ON, oculomotor nerve; OPN, optic nerve; PC, posterior commissure; PD, pars distalis of adenohypophysis; PEC, peduncle of cerebrum; PI, pineal body; PO, preoptic nucleus; PON, pons; PN, posterior nucleus; PT, pars tuberalis of adenohypophysis; PV, paraventricular nucleus; SO, supraoptic nucleus; SM, stria medullaris; VM, ventromedial nucleus.

from the hypothalamus. Each brain was cut midsagittally and each half of the hypothalamus was sectioned in the coronal plane to six strips (Fig. 1 A–F). In four brains one half of the hypothalamus was sectioned into 25 blocks (Fig. 1); the pituitary stalk was sectioned into seven horizontal slices and the infundibular stem, infundibular process (posterior pituitary) and the pars distalis (adenohypophysis) were also collected. Each of the sectioned pieces was weighed, extracted and analysed for GnRH and TRH by radioimmunoassay as described previously[6,7].

In a preliminary study we have found that in the rat the hypothalamic content of GnRH and TRH did not change up

to 48 h after death. Male rats (12 per group) were killed with ether. In one group, the hypothalami were removed and extracted immediately, and the other group kept for 48 h at 4 °C before removal and extraction of hypothalami. Hypothalamic content of GnRH was 4.4±0.3 ng at death compared with 4.0±0.3 ng 48 h later, and of TRH was 7.4±0.1 compared with 7.2±0.3 ng.

The content of GnRH and of TRH in the human hypothalamus were initially estimated by assaying the dissected strips (Fig. 1 A–F). The results are summarised in Table 1. The concentration of GnRH was the highest in the rostral strip
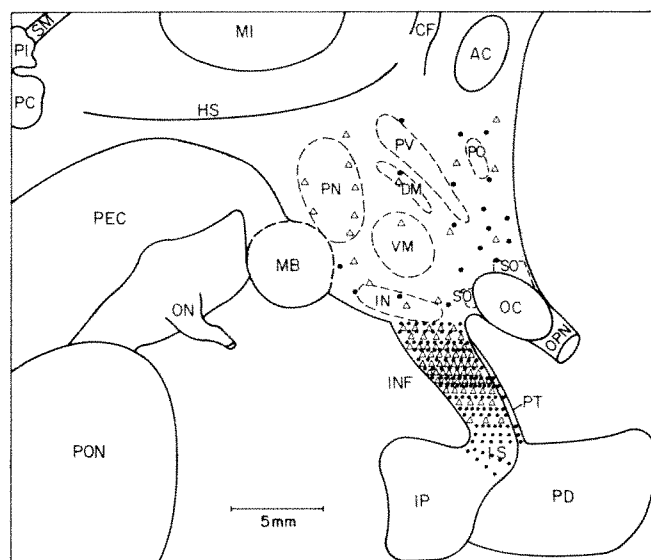
Fig. 2 Localisation of GnRH (●) and of TRH (△) in the human hypothalamus and pituitary stalk. The outline of hypothalamic nuclei in the sagittal plane are shown in broken lines. Each dot or triangle represents a concentration of 70–110 pg/mg tissue of the corresponding neurohormone. For abbreviations, see Fig. 1.

(A), which comprises the preoptic area, and declined progressively towards the caudal strip (F), which contained negligible amounts of GnRH. A different pattern of distribution was observed for TRH (Table 1): the highest concentration was found in strip C which consists of the posterior nucleus and the premammillary area. Radioimmunoassay of extracts derived from hypothalamic tissue blocks (Fig. 1), enabled us to follow more precisely the localisation of GnRH and TRH. GnRH was localised mainly in the preoptic area (Fig. 2), between the

Table 1 GnRH and TRH concentrations in the human brain

Sample	Weight (mg)	GnRH (pg per mg tissue)	TRH (pg per mg tissue)	Total GnRH (ng)	Total TRH (ng)
Hypothalamus					
Strip A	72 ± 14	130 ± 59	54 ± 10	9.36	3.88
B	98 ± 16	62 ± 24	73 ± 8	6.07	7.15
C	96 ± 5	27 ± 8.3	101 ± 14	2.59	9.69
D	105 ± 16	11 ± 1.6	69 ± 10	1.15	7.24
E	68 ± 5	12 ± 0.6	38 ± 12	0.81	2.58
F	54 ± 8	5.4 ± 3	5.2 ± 3	0.29	0.28
				20.27	30.82
Pituitary stalk*	41 ± 3	1,430 ± 334	855 ± 145	58.63	35.05
Infundibular process	82 ± 3	< 2	< 0.5		
Pars distalis	328 ± 48	< 2	< 0.5		
Pineal body	95 ± 20	< 2	1.1 ± 0.8		
Thalamus	91 ± 12	< 2	9.2 ± 2.6		
Cerebral cortex	100 ± 11	< 2	1.0 ± 0.4		

*Including infundibular stem.
Each tissue sample was weighed, extracted and analysed by radioimmunoassay for GnRH and TRH content. Mean ± s.e.m. for 10–15 samples, are shown. Total amount of the hormones in hypothalamic strips (A–F) represents the content of half a hypothalamus, dissected in the midsagittal plane.
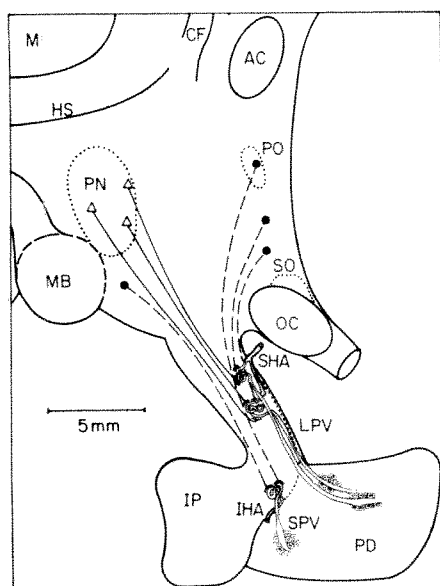
Fig. 3 Scheme of the neurovascular pathway of GnRH and TRH in the human brain. ●, Cell bodies of TRH-producing neurones in the posterior nucleus (PN). TRH is transferred by axons (solid lines) to the capillary loops in the pituitary stalk and reach the pars distalis (PD) of the adenohypophysis by way of the long portal vessels (LPV). ●, Cell bodies of GnRH-producing neurones in the preoptic area of the hypothalamus. Their axons (broken lines) terminate on the capillary loops of the superior hypophyseal artery (SHA) and the inferior hypophyseal artery (IHA), and reach the adenohypophysis by way of both the long (LPV) and short (SPV) portal vessels. The location of the portal vessels was drawn according to Daniel and Prichard[8]. For abbreviations, see Fig. 1.

supraoptic and the preoptic nuclei (340 pg per mg tissue); high concentration was also found in an area located anterior to the mammillary body in the basal hypothalamic area (186 pg per mg tissue). The highest concentration of TRH was observed in the posterior nucleus (158 pg per mg tissue) and in the region connecting this nucleus with the pituitary stalk. Exceedingly high concentrations of both neurohormones were present in the pituitary stalk (1,430 pg per mg tissue and 855 pg per mg tissue for GnRH and TRH, respectively; Table 1), with a difference in distribution (Fig. 2): GnRH was evenly distributed all along the pituitary stalk, whereas TRH was localised in the upper two-thirds (except for one case). The total content of GnRH in the hypothalamus was 40.5 ng and in the pituitary stalk 58.6 ng. Hypothalamic content of TRH was 61.7 ng and of the pituitary stalk 35 ng. No significant amounts of the hormones were detected in the different regions of the pituitary gland. Low concentrations of TRH were found in some extra-hypothalamic regions of the brain (Table 1). No apparent differences in the distribution or content of the two neurohormones were observed in relation to age or sex.

Axons descending from the hypothalamus have been shown to constitute the major component of the pituitary stalk[8]. Therefore, the extremely high concentrations of GnRH and TRH in the pituitary stalk (Fig. 2), may reflect the presence of pools of these hormones stored in axons and nerve terminals. Localisation of TRH and GnRH in specific hypothalamic regions may represent sites of synthesis of the hormones in neuronal cell bodies of these areas. In the rat, relatively high concentrations of GnRH have been found in the preoptic area and in the arcuate–median eminence region[9]. The median eminence of the rat, however, is anatomically homologous to the proximal portion of the human pituitary stalk[8]; therefore, the distributions of GnRH in human (that is, high concentrations in the preoptic area and the pituitary stalk) and rat hypothalamus, are similar.

An additional pathway—that is, by way of the cerebrospinal fluid—has been suggested[10] for the transfer of the neurohormones from their site of synthesis to their site of release. The high concentrations of GnRH in the lowest portion of the infundibular stem (Fig. 2), which is remote from the infundibular recess of the third ventricle, favours the direct neural route.

The absence of TRH in the lower third of the pituitary stalk implies that this neurohormone reaches the anterior pituitary by way of the long portal vessels (Fig. 3), whereas GnRH, which is evenly distributed along the whole pituitary stalk and infundibular stem, is transported by way of both the long and short portal vessels.

We thank Dr B. Bloch, Institute of Forensic Medicine, for assistance. E. O. is a visiting scientist from the Department of Pathology, Hadassah University Hospital, Jerusalem.

E. OKON
Y. KOCH

*Department of Hormone Research,
The Weizmann Institute of Science,
Rehovot, Israel*

1 Schally, A. V., Arimura, A., and Kastin, A. J., *Science*, 179, 341–350 (1973).
2 Guillemin, R., *Prog. Brain Res.* 41 117–132 (1974).
3 Harris, G. W., Reed, M., and Fawcett, C. P., *Br. med. Bull.*, 22, 266–272 (1966).
4 Schally, A. V., *et al.*, *J. clin. Endocr. Metab.*, 27, 755–762 (1967).
5 Winters, A. J., Eskay, R. L., and Porter, J. C., *J. clin. Endocr. Metab.*, 39, 960–963 (1974).
6 Koch, Y., Wilchek, M., Fridkin, M., Chobsieng, P., Zor, U., and Lindner, H. R., *Biochem. biophys. Res. Commun.*, 55, 616–622 (1973).
7 Koch, Y., Baram, T., and Fridkin, M., *FEBS Lett.*, 63, 295–298 (1976).
8 Daniel, P. M., and Prichard, M. M. L., *Acta Endocr. Copenh.*, suppl., 201, 1–216 (1975).
9 Wheaton, J. E., Krulich, L., and McCann, S. M., *Endocrinology*, 97, 30–38 (1975).
10 Knigge, K. M., and Scott, D. E., *Am. J. Anat.*, 129, 223–244 (1970).

Evidence for an extranucleolar mechanism of actinomycin D action

SINCE the original observation of Perry[1] that a low concentration of actinomycin D (0.04 µg ml⁻¹) selectively inhibits ribosomal RNA (rRNA) synthesis *in vivo*, this concentration of drug has been used by numerous investigators to selectively inhibit the synthesis of rRNA in cells in culture. Since nucleolar ribosomal DNA contains a high amount of deoxyguanosine and deoxycytosine bases[2] and actinomycin D selectively binds to deoxyguanosine[3], it has been thought that actinomycin D preferentially binds to nucleolar DNA to inhibit the transcription of 45S ribosomal RNA precursor.

α-Amanitin, the peptide toxin from *Amanita phalloides*, also has inhibitory effects on the synthesis of rRNA[4–6] and is known to produce nucleolar fragmentation[7–10]. Since α-amanitin is a potent inhibitor of nucleoplasmic RNA polymerase II (refs 11–13) and protein synthesis is an integral component of rRNA synthesis[14–17], it is conceivable that α-amanitin and actinomycin D have comparable sites of action in inhibiting the transcription of mRNA(s) coding for the essential protein(s) required for the initiation of rRNA transcription *in vivo*. Although these observations with α-amanitin have suggested the possibility of an extra-nucleolar control of rRNA synthesis, a specific role of mRNA in this process remains to be demonstrated. In this report, evidence is presented for an extranucleolar mechanism of actinomycin D action suggesting that rRNA transcription in nucleoli is controlled by mRNA synthesis in the nucleoplasm.

Studies in this report were designed to ask two questions: (1) Is α-amanitin resistant transcription in isolated nuclei (RNA polymerase I + III) and purified nucleoli (RNA polymerase I) inhibited by the low concentrations of actinomycin D known to be effective *in vivo*; and/or, (2)
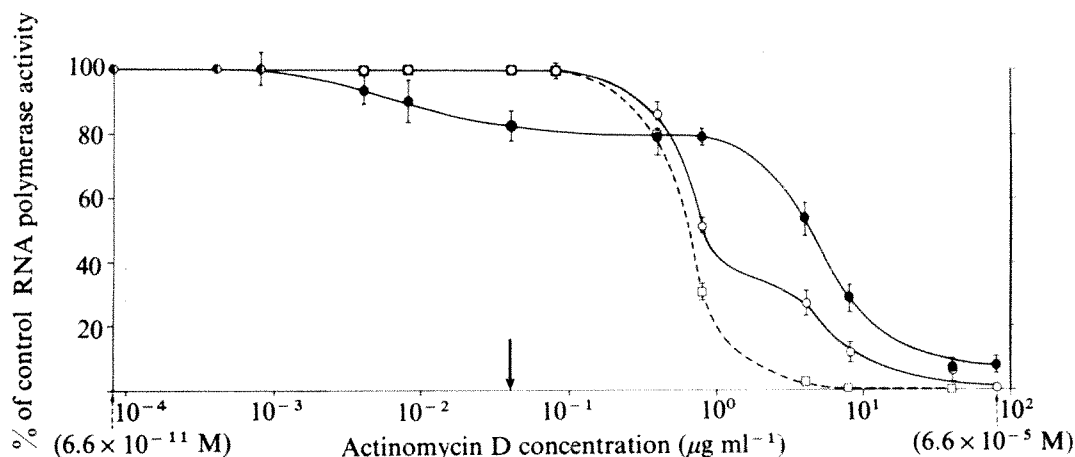
Fig. 1 Effect of actinomycin D on transcription in isolated nuclei and nucleoli. RNA polymerase assays, in the presence of varying concentrations of actinomycin D were performed as described below. Liver nuclei were isolated from adult male rats (300 g) by the method of Blobel and Potter[18] with 1 mM phenylmethylsulphonylfluoride (PMSF) in all buffers according to Lindell[19]. Nucleoli were isolated from freshly prepared nuclei by resuspension of the nuclear pellet obtained above in 0.34 M sucrose. The nuclear suspension was then sonicated and fractionated according to Roeder and Rutter[20]. The yield of nucleoli was 30–40% as calculated from the assay of RNA polymerase I activity in protein solubilised from starting nuclei and nucleolar product[21]. Routine preparations of nucleoli contain 0.5% RNA polymerase II by assay of RNA polymerase activity without exogenous DNA, plus and minus α-amanitin. RNA polymerase assays in rat liver nuclei and nucleoli were performed as described by Zerwekh *et al.*[22] using a single concentration of $(NH_4)_2SO_4$ (0.05 M). Actinomycin D (Merck) was added in 10-μl aliquots to 50-μl nuclear samples to give final concentrations (in 125 μl) ranging from 0.001 to 100 μg ml^{-1} (6.6 × 10^{-11}–6.6 × 10^{-5} M). Nuclei and nucleoli were preincubated with actinomycin D at room temperature for 10 min before initiation of the assays by addition of substrates and 5-^{3}H-UTP (New England Nuclear, 25 Ci mmol^{-1}). Individual nuclear assays (5 in the absence and 5 in the presence of α-amanitin, 0.1 μg per assay) were incubated for 5 min at 30 °C. Incorporation of ^{3}H-UMP was linear within this time period. Nucleolar assays were performed in the presence of α-amanitin to inhibit any possible residual RNA polymerase II activity. All incubations were terminated by pipetting 100-μl aliquots of each assay on to Whatman DE-81 filter disks and immediately placing them in 5% Na_2HPO_4. The filters were then washed, and counted as described by Lindell *et al.*[11]. Nucleoplasmic form II RNA polymerase activity was quantitated by subtracting the c.p.m. incorporated in the presence of α-amanitin from those assays performed in the absence of α-amanitin. Nuclear forms I and III RNA polymerase activities were quantitated by subtracting the c.p.m. obtained from control incubated samples containing 80 μg ml^{-1} actinomycin D plus α-amanitin (0.1 μg per assay) from the c.p.m. incorporated in the presence of α-amanitin alone. There is no linear incorporation of ^{3}H-UMP into RNA in this control. Activities of RNA polymerases are expressed as pmol UMP incorporated per 5 min per mg DNA. DNA determinations were performed on identical 50-μl aliquots of the same nuclear-nucleolar samples used for RNA polymerase assay by the method of Burton[23] using calf thymus DNA (Sigma, Type II) as a standard. Each point represents the average of five individual determinations ±s.e.m. Where error bars are not evident, they lie within the symbol. Control activities were: α-amanitin sensitive nucleoplasmic (form II RNA polymerase), 198; nuclear α-amanitin insensitive nucleoplasmic (forms I and III RNA polymerases), 114; and isolated nucleoli, 241 pmol UMP incorporated per 5 min per mg DNA, respectively. ●, Nucleoplasmic polymerase II; ○, nuclear RNA polymerases I and III (α-amanitin insensitive 0.1 μg per assay); and □, isolated nucleolar (RNA polymerase I) transcription. The arrow indicates 0.04 μg ml^{-1} actinomycin D.

Is there any nucleoplasmic component of *in vitro* transcription in nuclei which is inhibited by actinomycin D? Briefly, nuclei and nucleoli were isolated from adult rat livers and resuspended in a sucrose buffer. Actinomycin D was preincubated with these organelles before the addition of 5-^{3}H-UTP and substrates. Reactions were terminated by pipetting samples on to filter paper disks, washed, and counted as described in the legend to Fig. 1.

The effect of actinomycin D on the various RNA polymerase activities in isolated rat liver nuclei and purified nucleoli is shown in Fig. 1. Since actinomycin D inhibits rRNA transcription *in vivo* in the concentration range 0.001–0.1 μg ml^{-1} (ref. 24), it was surprising to observe that transcription in isolated nucleoli was inhibited at concentrations almost 100 times greater. Transcription in isolated nuclei in the presence of α-amanitin shows near identical inhibition to that observed in nucleoli except that the curve is biphasic. This additional activity between 1 and 10 μg ml^{-1} actinomycin D was inhibited by higher concentrations of α-amanitin (400 μg ml^{-1})[25] and, therefore, confirmed to be RNA polymerase III (data not included). An additional experiment was performed to rule out the possibility that actinomycin D was not preincubated with isolated nucleoli for an adequate amount of time for sufficient inhibition to occur. Figure 2 presents this experiment in which isolated nucleoli were incubated at 30 °C in the presence of 0.04 and 0.1 μg ml^{-1} actinomycin D relative to a control nucleolar sample incubated in the absence of actinomycin D. Samples of nucleoli were withdrawn at 15, 30, 45 and 60 min to test for the ability of the endogenous
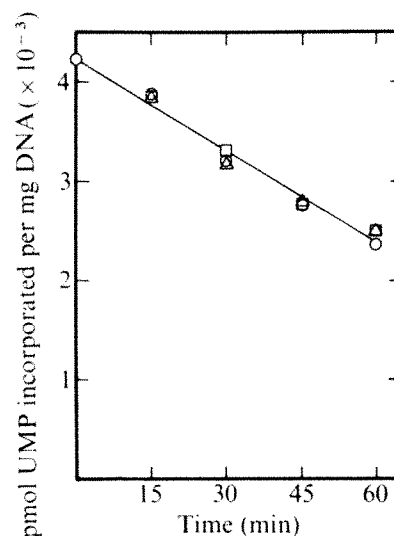
Fig. 2 Effect of actinomycin D preincubation with purified nucleoli on the activity of endogenous nucleolar RNA polymerase I. Actinomycin D at 0.04 and 0.1 μg ml^{-1} were preincubated with identical nucleolar samples at 30 °C for varying times up to 60 min. Aliquots (50 μl) were removed at the times indicated and pipetted into ^{3}H-UTP and substrates for assay of endogenous nucleolar polymerase activity as described in the legend of Fig. 1. Results obtained with actinomycin D are plotted along with an identical nucleolar sample which contained no actinomycin D. ○, Control; △, 0.04 μg ml^{-1} actinomycin D; and □, 0.1 μg ml^{-1} actinomycin D. RNA polymerase activity is expressed as pmol UMP incorporated per 10 min per mg DNA.

RNA polymerase I to transcribe nucleolar DNA. This experiment demonstrated that actinomycin D is still ineffective in inhibiting transcription in isolated nucleoli after pre-incubation of these organelles up to 1 h. The loss of activity with time may be related to the spontaneous loss of RNA polymerase I activity as previously described by Lindell[19] or to thermal inactivation of the enzyme.

Nucleoplasmic transcription by RNA polymerase II is biphasically inhibited by actinomycin D. Twenty per cent of the total nucleoplasmic RNA polymerase activity is inhibited by actinomycin D in the concentration range 0.001–0.1 $\mu g\ ml^{-1}$, whereas the remaining 80% is inhibited between 0.1 and 100 $\mu g\ ml^{-1}$. The first phase of nucleoplasmic polymerase transcription is inhibited at a concentration of actinomycin D 700 times lower than that required for the inhibition of the remaining nucleoplasmic activity.
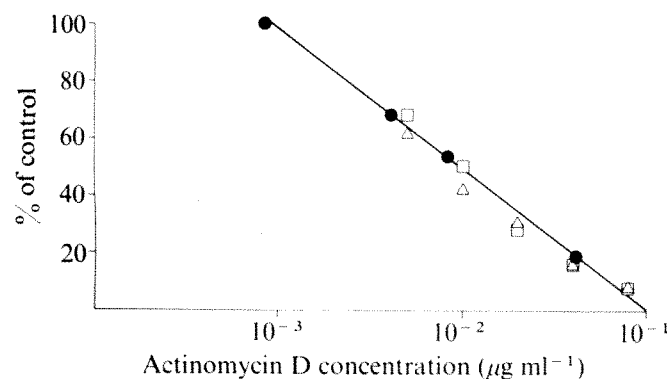
Fig. 3 Correlation of the inhibition of nucleoplasmic transcription *in vitro* with the inhibition of 45S RNA synthesis in mouse L cells *in vivo* (data of Perry and Kelley[24]) by actinomycin between 0.001 and 0.08 $\mu g\ ml^{-1}$. The first phase of nucleoplasmic transcription (Fig. 1) was relativised to 100% and compared with the data of Perry and Kelley[24] (their Figs 4 and 7, 100 and 60 min preincubation of mouse L cells with actinomycin D, respectively). ●, Nucleoplasmic transcription (this study); □, Perry and Kelley, 60-min preincubation (their Fig. 7); and △, 100-min preincubation (their Fig. 4).

Although actinomycin D inhibition of *in vitro* transcription seems to be reversed with regard to nucleolar and nucleoplasmic RNA synthesis, there is at least one consistency with another observation of actinomycin D inhibition in cells in culture relative to this work. It is clear from the work of Perry and Kelley[26] that 4S and 5S RNA synthesis which is transcribed by RNA polymerase III (ref. 25), is 50–100-fold less sensitive to inhibition than 45S RNA transcription in mouse L cells. This is also apparent in this *in vitro* work but the concentration range of actinomycin D which affects 45S RNA transcription *in vivo* only affects nucleoplasmic RNA synthesis (presumptive mRNA). The relationship of this actinomycin D inhibition of nucleoplasmic transcription in rat liver nuclei *in vitro* to the inhibition of 45S RNA synthesis in mouse L cells is shown in Fig. 3. It can be seen that an excellent correlation exists between these two heterologous observations.

Perhaps the most severe criticism of this work is that transcription in these organelles is not sufficiently sustained to allow any reasonable conclusions from a study such as this. Zylber and Penman[27] have concluded that transcription in Hela cell nuclei is merely completion of synthesis which had already initiated *in vivo*. In fact, transcription in isolated rat liver nuclei assayed in these conditions (polymerase II and polymerases I + III) is only linear for about 5 min. Transcription in isolated nucleoli assayed in identical conditions is, however, linear for 30–45 min. In spite of the lack of sustained transcription in isolated nuclei, there is a reasonable correlation with the inhibition of nucleolar transcription and the activity of α-amanitin-insensitive RNA polymerase activity in nuclei (Fig. 1).

The information presented in this manuscript lends further support to a possible extranucleolar control of rRNA transcription *in vivo*. Since actinomycin D only inhibits transcription by RNA polymerase II *in vitro* in the identical concentrations which inhibit 45S RNA transcription *in vivo*, it is likely that this extranucleolar control resides at the level of mRNA transcription. In fact, other investigators have observed an inhibition of nucleoplasmic RNA transcription by low concentrations of actinomycin D in eukaryotic cells in culture[24,27,29]. The amount of nucleoplasmic RNA inhibited ranged from 5 to 50% depending on the cell type and the concentration of actinomycin D used.

If rRNA transcription is controlled by nucleoplasmic RNA polymerase II transcription (mRNA) as suggested by these data, then a rational hypothesis can be formulated with regard to the early observations of Fiume and Laschi[7] regarding nucleolar fragmentation induced by α-amanitin. While inhibitory effects of α-amantin on rRNA synthesis have been observed[4–6], others have seen no effect of this compound in specific cell types[10,30]. It is, however, clear from these studies that α-amanitin probably does not easily penetrate the cell membrane because nucleolar fragmentation is not evident until 6–8 h after addition of α-amanitin (5 $\mu g\ ml^{-1}$) to CHO cells and effects on rRNA synthesis are not observed until 24 h (ref. 10). For α-amanitin to be an effective inhibitor of rRNA transcription by the proposed pathway, all RNA polymerase II molecules must be inhibited to be assured that the expression of any specific genome is terminated. Actinomycin D, on the other hand, probably acts at a specific DNA site(s) which has a very high affinity for binding and, hence, a selective inhibitory action. These data would predict that that site is nucleoplasmic rather than nucleolar.

Thomas J. Lindell

Department of Pharmacology,
University of Arizona College of Medicine,
Tucson, Arizona 85724

1 Perry, R. P., *Expl Cell Res.*, 29, 400–406 (1963).
2 Birnsteil, M. L., Spiers, J., Purdom, I., Jones, K., and Loening, U. E., *Nature*, 219, 454–463 (1968).
3 Sobell, H. M., Jain, S. C., Sakore, T. D., and Nordman, C. E., *Nature new Biol.*, 231, 200–205 (1971).
4 Jacob, S. T., Muecke, W., Sajdel, E. M., and Munro, H. N., *Biochem. biophys. Res. Commun.*, 40, 334–342 (1970).
5 Tata, J. R., Hamilton, M. J., and Shields, D., *Nature new Biol.*, 238, 161–164 (1972).
6 Sekeris, C. E., and Schmid, W., *FEBS Lett.*, 27, 41–45 (1972).
7 Fiume, L., and Laschi, R., *Sper. Arch. Biol. Norm. Path.*, 115, 288–297 (1965).
8 Marinozzi, V., and Fiume, L., *Expl Cell Res.*, 67, 311–322 (1971).
9 Petrov, P., and Sekeris, C. E., *Expl Cell Res.*, 69, 393–401 (1971).
10 Kedinger, C., and Simard, R., *J. Cell Biol.*, 63, 831–842 (1974).
11 Lindell, T. J., Weinberg, F., Morris, P. W., Roeder, R. G., and Rutter, W. J., *Science*, 170, 447–449 (1970).
12 Kedinger, C., Gniazdowski, M., Mandel, J. L., Gissinger, F., and Chambon, P., *Biochem. biophys. Res. Commun.*, 38, 165–171 (1970).
13 Jacob, S. T., Sajdel, E., and Munro, H. N., *Nature*, 225, 60–62 (1970).
14 Willems, M., Penman, M., and Penman, S., *J. Cell Biol.*, 41, 177–187 (1969).
15 Muramatsu, M., Shimada, N., and Higashinakagawa, T., *J. molec. Biol.*, 53, 91–106 (1970).
16 Craig, N. C., and Perry, R. P., *J. Cell Biol.*, 45, 554–564 (1970).
17 Yu, F. L., and Feigelson, P., *Proc. natn. Acad. Sci. U.S.A.*, 68, 2177–2180 (1971).
18 Blobel, G., and Potter, V. R., *Science*, 154, 1662–1665 (1966).
19 Lindell, T. J., *Archs Biochem. Biophys.*, 171, 268–275 (1975).
20 Roeder, R. G., and Rutter, W. J., *Proc. natn. Acad. Sci. U.S.A.*, 65, 675–682 (1970).
21 Gamo, S., and Lindell, T. J., *Life Sci.*, 15, 2179–2187 (1974).
22 Zerwekh, J. E., Haussler, M. R., and Lindell, T. J., *Proc. natn. Acad. Sci. U.S.A.*, 71, 2337–2341 (1974).
23 Burton, K., *Biochem.*, 62, 315–323 (1956).
24 Perry, R. P., and Kelley, D. E., *J. Cell Physiol.*, 76, 127–140 (1970).

[25] Weinmann, R., and Roeder, R. G., *Proc. natn. Acad. Sci. U.S.A.*, 71, 1490–1494 (1974).
[26] Perry, R. P., and Kelley, D. W., *J. Cell Physiol.*, 72, 235–246 (1968).
[27] Zylber, E. A., and Penman, S., *Proc. natn. Acad. Sci. U.S.A.*, 68, 2861–2865 (1971).
[28] Penman, S., Vesco, C., and Penman, M., *J. molec. Biol.*, 34, 49–69 (1968).
[29] Lindberg, U., and Persson, T., *Eur. J. Biochem.*, 31, 246–254 (1972).
[30] Chambon, P., *A. Rev. Biochem.*, 44, 613–638 (1975).
[31] Simard, R., *J. Cell Biol.*, 35, 716–722 (1967).
[32] Halbers, E., Muller, W., and Backmann, R., *Biochem. Z.*, 337, 224–231 (1963).
[33] Goldstein, M. N., Hamm, K., and Amrod, E., *Science*, 151, 1555–1556 (1966).

Possible pathway for prebiotic uracil synthesis by photodehydrogenation

THE synthesis of purines in potentially prebiotic conditions has been demonstrated[1,2]. Adenine and guanine have been synthesised directly from aqueous cyanide solutions and a mechanism for the concentration of hydrogen cyanide in low melting eutectics has been proposed[3]. These purines are also obtained from solutions of cyanogen (unpublished). The pyrimidines, however, have provided a more difficult problem. Uracil has been obtained from malic acid and urea, but only by heating in the presence of concentrated sulphuric acid[4] or polyphosphoric acid[5]. Low yields of uracil have been inferred from spectral data for a reaction involving acrylonitrile, urea and ammonium chloride[6]. Further details concerning the conditions of this reaction, however, have not been offered. Cytosine and, indirectly uracil, have been prepared from cyanoacetylene and cyanate[7,8]. The rapid hydrolysis of these reactants raises doubts about this route. Cyanoacetaldehyde, a hydrolysis product of cyanoacetylene, condenses with guanidine in aqueous solution to yield 2,4-diaminopyrimidine, which is hydrolysed to cytosine and uracil. The later route is considered to be more plausible[9]. Thymine has been prepared in 0.1% yield from uracil by a hydroxymethylation–reduction sequence involving formaldehyde and hydrazine hydrate in ammoniacal solution[10]. In view of uncertainties concerning the abundances on the primitive Earth of a number of the precursors which have been used in these studies, it seemed desirable to consider alternative routes to the synthesis of pyrimidines. In this report we describe the synthesis of uracil by the photodehydrogenation of 5,6-dihydrouracil (DHU), as well as the synthesis of DHU from β-alanine and urea in mild conditions.

β-Ureidopropionic acid (UPA), which is readily formed from β-alanine by reaction with urea[11] or with cyanate[12,13], is known to cyclise to DHU. The cyclisation is catalysed by acid[12], although certain substituted derivatives have been cyclised by heat[14]. We have found that the cyclisation of UPA can be affected by heating at moderate temperatures. UPA (2 g) in 20 ml H₂O was added to 25 g silica gel (Merck Kieselgel 60) and the water evaporated *in vacuo*. The mixture was then heated at 65 °C under N₂ for 5 weeks. DHU was isolated by column chromatography on cellulose in butan-2-ol–water (86 : 14) and characterised as *N*-acetyl-5,6-dihydrouracil (207 mg) by mixed melting point and infrared spectroscopy[15]. We have observed that the cyclisation of UPA to form DHU is promoted by urea. The direct evaporation of dilute solutions of β-alanine and urea either alone or in the presence of various silicates also results in the formation of DHU in the temperature range 60–90 °C. Further details of these experiments will be reported elsewhere.

In a typical photodehydrogenation experiment, 50 mg DHU in 5 ml hot distilled water was added to 1 g support material (Table 1), contained in a borosilicate tube (160 × 50 mm). The tube was equipped with inlet and outlet side arms for N₂ purging and a ground-glass joint for insertion of a Hanovia low-pressure mercury arc lamp. The tube was fitted to a rotary evaporator and the slurry evaporated and deposited on the walls of the tube *in vacuo*. Irradiation was carried out under N₂ for 48 h. The effect of water vapour on the reaction was tested by passing the N₂ through a flask filled with distilled water or, alternatively, through KOH. At the end of the irradiation period, the material was extracted with 100 ml hot distilled water and filtered. The products from these experiments gave
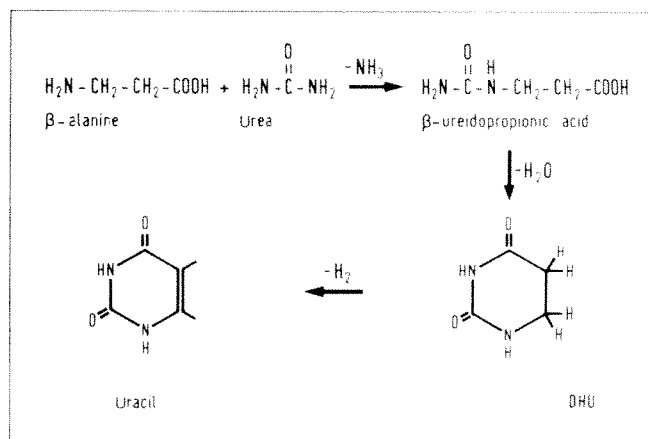
Fig. 1 Postulated pathway for prebiotic formation of uracil.

positive Wheeler–Johnson tests[16]. The presence of uracil was confirmed by ultraviolet absorption spectroscopy and by thin-layer chromatography on silica gel (*n*-butanol–acetic acid–water; 3 : 3 : 1) and on cellulose (*tert*-butanol–methyl ethyl ketone–water–concentrated NH₃, 40 : 30 : 20 : 10). Products from a number of experiments were pooled, and uracil was isolated by preparative thin-layer chromatography on cellulose as described above. Acetylation gave *N*-acetyluracil, as shown by comparison with an authentic sample by infrared spectroscopy and mixed melting point[15]. Quantitative analysis of uracil was carried out by ion-exclusion chromatography on Aminex A25 and A6 columns as described previously[17,18].

Results are summarised in Table 1. In the absence of added support, 2% conversion of DHU to uracil was obtained after

Table 1 Conversion of 5,6-dihydrouracil to uracil on various supports

Support*	Yield of uracil (mole %)†
Montmorillonite number 20	22 (6)
Montmorillonite number 23	14 (7)
Montmorillonite number 31	12 (5)
Illite number 35	5 (2)
Illite number 36	4
Vermiculite (Zonolite)	3
Kaolinite number 9	2
Kaolinite number 5	1
Silica gel	3 (2)
Quartz	2
Basalt	2
Granite	1
CaCO₃	2
None	2 (1)

*CaCO₃ and silica gel (Kieselgel H) were from Merck. Clay minerals were standard samples from Ward's and were hand ground. Quartz sand (Baker) and the basalt and granite samples were ground in a disk mill (Tema).
†Irradiation for 48 h under water-saturated N₂ as described in the text. Figures in parentheses are yields obtained with dry N₂.

48 h. In the presence of montmorillonite number 20, however, the yield of uracil was 22%. This difference, as well as the influence of water vapour on the yields, suggests that the most active materials exert an effect by binding water in the solid phase and perhaps catalysing its photolytic dissociation. (Irradiation of DHU in aqueous solution also yielded uracil. After 4 d a 7% yield was obtained from a 0.0044 M solution.) Destruction of starting material and/or product was also more severe in the presence of water vapour. Thus, in the experiments with montmorillonite, uracil + DHU accounted for 90 to 97% of the starting material in the absence of water vapour, but only 74 to 82% in its presence. The series—montmorillonite, illite, vermiculite, kaolinite—is particularly interesting, as it represents a transition from expandable lattice structure

(montmorillonite) to non-expandable (kaolinite)[19]. A plausible mechanism for the photodehydrogenation might involve dissociation of bound water molecules and the hydroxylation and subsequent dehydration of DHU. The mechanism may therefore be similar to that involved in the ⁶⁰Co γ radiolysis of DHU, in which 5-hydroxy-, 6-hydroxy- and 5,6-dihydroxy DHU are formed[20]. Studies on the γ radiolysis of dihydrothymine have also implicated the OH radical as the reactive species in the formation of thymine[21]. Our studies on the mechanism of the catalysed photodehydrogenation are continuing.

We suggest that the synthesis of uracil by the reaction sequence shown in Fig. 1 represents a reasonable prebiotic pathway. Urea has been repeatedly suggested as a prebiotic reagent[22]. As the product formed in highest yield from solutions of hydrogen cyanide as well as by other pathways, it is one of the most plausible nitrogen-containing prebiotic materials and has been proposed as a possible storage form for ammonia on the primitive Earth[23]. β-Alanine has been reported in the Murchison meteorite[24], and has been formed in a number of simulated prebiotic syntheses[25-27]. Its abundance relative to other amino acids varies depending on the conditions chosen. In our proposed pathway, β-alanine would probably have been limiting. The entire pathway of carbamylation, cyclisation and photodehydrogenation may have occurred in evaporating pond environments, particularly where cycles of dehydration and rehydration were common.

We thank Matty Goverde for assistance.

G. J. F. CHITTENDEN
ALAN W. SCHWARTZ

Department of Exobiology,
University of Nijmegen,
Toernooiveld, Nijmegen, The Netherlands

1 Oró, J., and Kimball, A. P., *Archs Biochem. Biophys.*, 94, 217–227 (1961)
2 Sanchez, R. A., Ferris, J. P., and Orgel, L. E., *J molec. Biol*, 30, 223–253 (1967) (and refs therein)
3 Sanchez, R. A., Ferris, J. P., and Orgel, L. E., *Science*, 153, 72–73 (1966).
4 Davidson, D., and Baudisch, O., *J. Am chem. Soc*, 48, 2379–2383 (1926)
5 Fox, S. W., and Harada, K., *Science*, 133, 1923–1924 (1961)
6 Oró, J., *Fedn Proc.*, 22, 681 (1963).
7 Sanchez, R A., Ferris, J P., and Orgel, L E., *Science*, 154, 784–785 (1966)
8 Ferris, J P., Sanchez, R A., and Orgel, L E., *J molec Biol*, 33, 693–704 (1968).
9 Ferris, J P., Zamek, O. S., Altbuch, A. M, and Freiman, H, *J. molec. Evol*, 3, 301–309 (1974).
10 Stephen-Sherwood, E., Oró, J, and Kimball, A P., *Science*, 173, 446–447 (1971).
11 Br Pat 913, 713 (1962), *Chem Abs*, 58, 12428h (1963), Swiss Pat. 380, 716 (1965), *Chem. Abs*, 62, 16069f (1965).
12 Langfeld, F, and Stoeglitz, J J., *Am. chem J*, 15, 504–518 (1893)
13 Fink, R. M., McGaughey, C, Chne, B E., and Fink, K, *J biol. Chem.*, 218, 1–7 (1956).
14 Posner, T, and Rohde, K., *Ber.*, 42, 2785–2794 (1909)
15 Spector, L B., and Keller, E B, *J biol Chem*, 232, 185–192 (1958)
16 Wheeler, H. L., and Johnson, T. B., *J. biol. Chem.*, 3, 183–189 (1907).
17 van der Velden, W, and Schwartz, A. W., *Science*, 185, 691–693 (1974).
18 van der Velden, W., Chittenden, G. J F., and Schwartz, A W., *Advances in Organic Geochemistry*, 1973 (edit by Tissot, B., and Bienner, F.), 293–304 (Editions Technip, Paris, 1974)
19 Grimm, R E, *Clay Mineralogy*, 49 (McGraw-Hill, New York, 1968)
20 Vincendon, G, Crer, A, and Nofre, C, *Bull Soc. chim France*, 32, 1997–2000 (1965)
21 Snipes, W., and Bernhard, W, *Radiat Res*, 33, 162–173 (1968)
22 Lohrmann, R, and Orgel, L E, *Science*, 171, 490–494 (1971) (and refs therein).
23 Ferris, J P., Nicodem, and D. E., in *The Origin of Life and Evolutionary Biochemistry* (edit by Dose, K., Fox, S W, Deborin, G. A., and Pavlovskaya, T. E.), 107 (Plenum, New York, 1974)
24 Kvenvolden, K. A., Lawless, J. G., and Ponnamperuma, C, *Proc natn. Acad. Sci. U.S.A*, 68, 486–490 (1971)
25 Miller, S L, *Biochim biophys. Acta*, 23, 480–487 (1957), Oró, J, Kimball, A., Fritz, R, and Master, F, *Archs Biochem Biophys*, 85 115–130 (1959)
26 Wolman, Y, Haverland, W J., and Miller, S. L., *Proc. natn Acad. Sci U.S.A*, 69, 809–811 (1972).
27 Park, W. K., Hochstim, A. R., and Ponnamperuma, C, *Origins of Life*, 6, 99–107 (1975).

Effect of point-freezing on ethylene and ethane production by sugar beet leaf disks

THE site of biosynthesis of the plant hormone ethylene in the cells of higher plants is unknown. Efforts to identify the enzyme system responsible and its subcellular location have failed because ethylene-producing plant tissues stop pro-

duction after homogenisation. On the other hand, ethane is produced as the main hydrocarbon gas after homogenisation of certain tissues of higher plants[1,2]. We have investigated the influence of point freezing on ethylene and ethane synthesis by sugar beet leaf disks, and found that both are stimulated.

Rates of ethylene and ethane formation were monitored after incubation of untreated or point frozen leaf disks of 23 mm diameter and about 100 mg fresh weight, in order to show that sugar beet leaf tissue behaves as theoretically expected[1,2]. The disks were cut from the upper third of sugar beet leaves not more than 12 weeks old, and floated on 1 ml of water in 3-ml Fernbach flasks, closed by a serum rubber stopper. After 2 h of incubation at 25 °C in the dark, 0.5 ml of the gas phase was analysed for ethylene and ethane by gas chomatography (compared with authentic gases purchased by Messer-Griesheim GmbH, Dusseldorf) as described before[3].

Table 1 Ethylene and ethane production of untreated and pressed sugar beet leaf disks

	Ethylene (a)	Ethane (b)	Ratio a/b
	(pmol per h per g fresh weight)		
Untreated disks	379±355 (61)	1±3 (35)	95±320 (35)
Pressed disks	33±171 (41)	86±40 (60)	0.4±2.6 (41)

Results are mean ±s.d. Figures in parentheses are numbers of experiments. These data were compared by statistical methods (t test, F test). Rate of ethylene production of untreated disks > pressed disks (P<0.01). Rate of ethane production of pressed disks > untreated disks (P<0.01). Ratio a/b of untreated disks > pressed disks (P<0.01).

The most ethylene was formed by untreated leaf disks (Table 1). Production of ethylene was strongly reduced after the leaf disks had been pressed with a plastic piston on a plastic board. In contrast, ethane production was low in untreated leaf disks and was strongly stimulated after pressing. These results, obtained by pressing sugar beet leaf disks, are in good agreement with results obtained after homogenisation of various higher plant tissues[1,2].

The following experiments demonstrate a correlation between the extent of wounding and the rates of production of ethylene and ethane by sugar beet leaf disks. For this purpose, leaf disks of 5-week-old sugar beets were frozen to various extents: a stainless steel rod, 3 mm in diameter, kept at the temperature of liquid nitrogen, was touched to the surface of the leaf disks for 2 s. This resulted in a frozen circular area of approximately 3% of the total leaf disk area. By repeating this process, leaf disks were point frozen to 25, 50, 75 and 100% with respect to the total leaf disk area and were then incubated as described above. Figure 1 shows the rates of ethylene and ethane production plotted against the percentage of frozen leaf disk areas. Production of ethylene increased as the extent of point freezing of the total leaf disk area increased from 0 to 50%, and decreased linearly as the frozen leaf disk area increased from 50 to 100%.

In contrast to ethylene, ethane formation increased linearly as the point freezing of the total leaf disk area increased from 0 to 100% (Fig. 1). The measured ethylene to ethane ratios plotted against the percentage of frozen leaf disk area (Fig. 2) shows a strict logarithmic dependence.

The linear stimulation of ethane production after leaf disk point freezing up to 100% of the disk area indicates that, in contrast to ethylene, ethane production is not dependent on intact compartmentalisation in the leaf cells, but rather depends on cellular disorder. Thus, ethane production and the ethylene to ethane ratio can be taken as an indicator of the integrity of sugar beet leaf tissue.
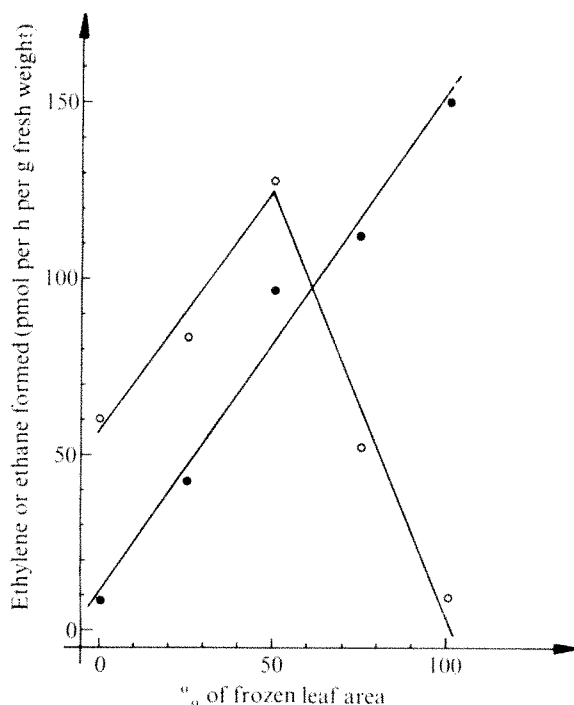
Fig. 1 Dependence of ethylene (○) and ethane (●) production on the percentage of frozen leaf area. Point freezing and incubation were achieved as described in the text. The points of the lines represent the mean of five experiments.

The influence of freezing on ethylene production seems to be more complex. Up to a certain extent of point freezing (in our case 50% of frozen leaf disk area), ethylene production increased after point freezing in spite of the decrease of unfrozen leaf area. When 50% of leaf disk area was frozen, further point freezing severely decreased ethylene production to an almost negligible amount. Assuming a K_m for O_2 of 0.2% for ethylene synthesis[4], the decrease of ethylene production and the concomitant increase of ethane production cannot be attributed to oxygen depletion as the result of an increased oxygen consumption after point freezing. This can be concluded from the observation that pressed leaf disks of 100 mg fresh weight consume about 1.6 μmol of O_2 per 2 h,
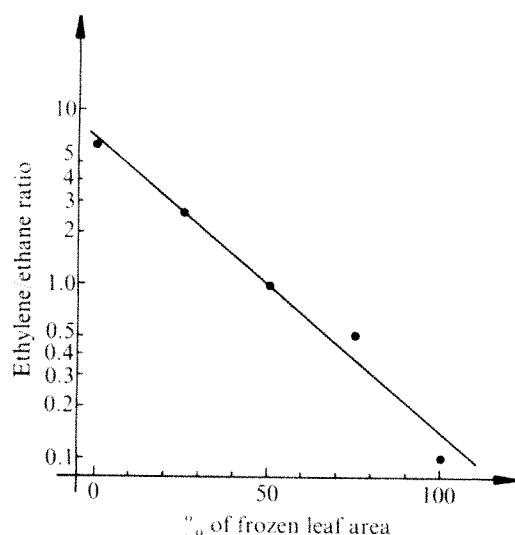
Fig. 2 Dependence of the ethylene to ethane ratio on the percentage of frozen leaf area. Point freezing and incubation were achieved as described in the text. The points of the lines represent the mean of five experiments.

representing 8% of the total amount (20.0 μmol) of O_2 available. From the above results we conclude that ethylene formation—in contrast to ethane formation—occurs in the leaf area surrounding the frozen parts of the leaf, representing a wound response of the non-decompartmentalised, but physiologically perturbed[5,6] cells, adjacent to the decompartmentalised (killed) cells.

If the percentage of the decompartmentalised tissue was increased, the non-decompartmentalised, but perturbed tissue was reduced, this being observed as a decrease in ethylene production.

The physiological role of the stimulation of ethylene production after wounding can be seen in the induction of wound responses, for example, the accumulation of phenolic compounds[7]. It is thus interesting that injury-related biosynthesis of an aromatic phytoalexin (phaseollin) in *Phaseolus vulgaris* has also been described to be dependent on living tissue adjacent to the injuries[8]. Ethane production, on the other hand, does not depend on intact tissue but rather on cellular disorder and loss of compartmentalisation, for the highest production rates are obtained with 100% frozen leaf area. Because of this independent behaviour of ethylene and ethane production, two different pathways can be assumed for their biosynthesis.

E. F. ELSTNER
J. R. KONZE

Department of Plant Biochemistry,
Ruhr University Bochum,
PO Box 2148,
4630 Bochum, FRG

1 Lieberman, M., and Mapson, L. W., *Nature*, **195**, 1016–1017 (1962).
2 Curtis. R. W., *Pl. Physiol., Lancaster*, **44**, 1368–1370 (1969).
3 Elstner, E. F., and Konze, J. R., *Z. Naturforsch.*, **29c**, 710–716 (1974).
4 Burg, S. P., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 591–597 (1973).
5 Irvine, R. F., and Osborne, D. J., *Biochem. J.*, **136**, 1133–1135 (1973).
6 Hiraki, Y., and Ota, Y., *Pl. Cell Physiol.*, **16**, 185–189 (1975).
7 Hislop, E. L., Hoad, G. V., and Archer, S. A., in *Fungal Pathogenicity and the Plant's Response* (edit. by Byrde, R. J. W., and Cutting, C. V.), 87–117 (Academic, London and New York, 1973).
8 Rahe, J. E., and Arnold, R. M., *Can. J. Bot.*, **53**, 921–928 (1975).

Errata

In the article "Significance of impulse activity in the transformation of skeletal muscle type" by S. Salmons and F. A. Sréter (*Nature*, **263**, 30; 1976), the tenth line from the bottom of the first column on page 33 should read . . . soleus muscles which had received stimulation showed light- . . . and not as printed.

In the article "Opposing effects of cyclic AMP and cyclic GMP on protein phosphorylation in tubulin preparations" by I. V. Sandoval and P. Cuatrecasas (*Nature*, **262**, 511; 1976) the pH value given in the second line of the legend to Table 1 should read pH 6.75 and not 7.65 as printed.

matters arising

Green rust: a pyroaurite type structure

McGILL, McENANEY AND SMITH[1] have described a green corrosion product of iron developed in water containing NaCl, K_2SO_4, $NaCO_3$, mixtures of these salts, and NaOH under various degrees of aeration. The X-ray data they reported indicate that the product is a member of the pyroaurite group[2-4]. In fact, Stampfl[5] and Taylor[6] have also concluded that a ferrous-ferric green rust material belongs to the group of pyroaurite structures.

We have been concerned[7] with the synthesis of products similar to the mineral takovite which has a formula of the pyroaurite type near to $(Ni,Mg)_6$ $(Al,Fe^{3+})_2(OH)_{16}CO_34H_2O$. We find that such products are formed readily in solutions containing Cl^-, NO_3^- and SO_4^{2-} ions, provided that CO_2 is not rigorously excluded. When water in its normal state or saturated with CO_2 is used, the same or similar final product is obtained. The presence of the anion CO_3^{2-} is shown unambiguously by infrared spectroscopic data. We suspect that other anions play only a small role in the formation of the product when CO_3^{2-} ions are present. For reasons which are not entirely clear, the pyroaurite structure type seems to form very readily when CO_3^{2-} ions are available. Possibly the planar nature of the CO_3^{2-} ion and the total number required are favourable factors; NO_3^- ions, although planar, must be twice as numerous for electrical neutrality.

Using the X-ray data reported by McGill, *et al.*[1], we have recalculated the unit cell parameters using the refinement program of Evans *et al.*[7]; the results are: $a = 3.166 \pm 0.001$, and $c = 22.52 \pm 0.01$ Å —values which are close to those given by McGill *et al.* and also by Stampfl[5]: $a = 3.17$; $c = 22.8$ Å.

It is of interest to consider a possible relationship between cation composition (as indicated by the weighted mean cation radius, $\bar{r}$), and the hexagonal a parameter. Figure 1 shows such a plot with data taken from the literature, including two unpublished points (G.W.B. and D.L.B.), dealing with takovite. The $\bar{r}$ values are calculated using the crystal radii of Shannon and Prewitt[8]. The smallest value of $\bar{r}$ is for takovite (G.W.B. and D.L.B., unpublished) and the largest for the green rust studied by Butler and Benyon[9] on the assumption of a 3 : 1 ferrous-ferric ratio. The available data cluster near to a
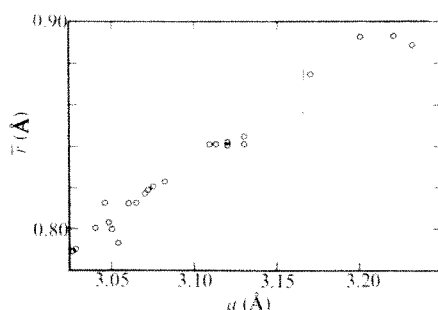
Fig. 1 Plot of $\bar{r}$ (mean cation radius), against unit cell parameter a for pyroaurite-type structures. Vertical line indicates the a parameter of the green rust of McGill *et al.*[1].

straight line and indicate for the green rust of McGill *et al.* a ferrous-ferric ratio of about 2:1; the precise result depends on how the line is drawn. In future studies of pyroaurite group materials, it will be useful to have chemically analysed materials and accurate cell parameters so that a composition-cell parameter graph can be drawn with greater accuracy than is now possible.

G. W. BRINDLEY
DAVID L. BISH

126 Mineral Sciences Building,
Department of Geosciences,
The Pennsylvania State University,
University Park, Pennsylvania 16802

1 McGill, I. R., McEnaney, B., and Smith, D. C., *Nature*, **259**, 200–201 (1976).
2 *X-Ray Powder Data File*, Card No. 14.293 (Joint Committee on Powder Diffraction Standards, 1601 Park Lane, Swarthmore, Pennsylvania, 19081).
3 Mumpton, F. A., Jaffe, H. W., and Thompson, C. S., *Am. Miner.*, **50**, 1893–1913 (1965).
4 Allmann, R., *Chimia*, **24**, 99–108 (1970).
5 Stampfl, P. P., *Corros. Sci.*, **9**, 185–187 (1969).
6 Taylor, H. F. W., *Mineralog. Mag.*, **39**, 377–389 (1973).
7 Evans, H. T., Appleman, D. E., and Handwerker, D. S., *Proc. Am. crystallogr. Ass.*, Annual Meeting Program, 42–43 (1963).
8 Shannon, R. D., and Prewitt, C. T., *Acta crystallogr.*, **B25**, 925–946 (1969).
9 Butler, G., and Benyon, J. G., *Corros. Sci.*, **7**, 385–404 (1967).

AN extraordinary orientational relationship between the rhombohedral green rust (with hexagonal unit cell dimensions $a = 3.181$ Å and $c = 21.82$ Å) and magnetite (cubic with $a = 8.434$ Å)— namely, $[1\bar{1}1]_{rh}$ or $[42.\bar{1}]_{hex}$ // $[1\bar{1}1]_{cub}$; and $(566)_{rh}$ or $(1\ 0\ \bar{1}\ 17)_{hex}$ // $(1\bar{1}1)_{cub}$—has been reported by McGill, McEnaney and Smith[1]. I should like to query this result, and to suggest that there is really no evidence that the relationship is other than $[00.1]_{hex}$ or $[111]_{rh}$ // 111_{cub}, and $(30\bar{3}0)_{hex}$

or $(2\bar{1}\bar{1})_{rh}$ // $(112)_{cub}$, as would be expected from a comparison of the oxygen packings of the two structures.

If the material is not well crystallised, or the platelet is very thin, the reciprocal lattice points will be extended parallel to $[00.1]_{hex}$ and it will not then be necessary to postulate that the diffraction spots, visible in their Fig. 1 and obtained when the beam is perpendicular to the crystal flake, have indices which strictly obey the requirement $-h + k + l = 3n$. The photographs reproduced in ref. 1 do not seem to show departures from hexagonal symmetry. Moreover, the hexagonal and rhombohedral cell dimensions are not entirely consistent with each other and the sets of planes considered to be equivalent do not really have similar d-spacings.

The simpler topotactic relationship had been reported earlier[2,3] but, together with work on the structure of the green rust materials[4], had gained inadequate publicity.

A. L. MACKAY

Department of Crystallography,
Birkbeck College, (University of London),
Malet Street, London WC1E 7HX, UK

1 McGill, I. R., McEnaney, B., and Smith, D. C., *Nature*, **259**, 200–201 (1976).
2 Mackay, A. L., in *Reactivity of Solids*, (edit. by de Boer, J. H.), *Mineralog. Mag.*, **33**, xlviii (1962); **36**, xxxiv (1968); 571–583 (Amsterdam, 1960).
3 Lea, S. G., thesis, Univ. London (1968).

McGILL ET AL. REPLY—Mackay's argument[1] presumes that there is close matching of oxygen ions in the green rust and magnetite structures resulting from a topotactic transformation, that is, internal atomic displacements resulting in accord in three dimensions between initial and final lattices[2]. There is no conclusive evidence for this type of transformation from our experiments on the corrosion of cast iron. It has been found that magnetite and green rust usually precipitate from solution independently of each other, although there are isolated examples of the epitaxial growth of magnetite on green rust from solution. Figure 3 of our original contribution[3] illustrates this, and shows magnetite as three-dimensional, 'island' outgrowths from the surface of green rust crystals, a morphology commonly observed in epitaxy. Close matching of oxygen ions between green rust and magnetite is not, therefore, a necessary condition, since epitaxy can occur with large degrees of atomic

mismatch between substrate and overgrowth[4].

The orientation relationship that we report results from a correlation of *d*-spacings from electron diffraction, and the more accurate *d*-spacings from X-ray diffraction; such a correlation cannot be obtained if the orientation relationship proposed by Mackay[1] is assumed. Nor is there evidence to postulate a relaxation in the Laue condition, since scanning electron microscopy (see Fig. 3, ref. 3) indicates that the green rust crystals are ~ 0.1 μm in thickness or greater. The elongation of diffraction spots is usually associated with microcrystals which have critical dimensions of much less than 0.1 μm. Electron diffraction suggests that the green rust platelets are well-crystallised. An inspection of the rhombohedral unit cell shows that a flat crystal with a $(5\ \bar{6}\ \bar{6})$ habit plane may have hexagonal geometry.

Turning to the point raised by Brindley and Bish[5]: in our experiments green rust is grown in closed vessels by the corrosion of cast iron in solutions which are deoxygenated by purging with nitrogen. Thus the level of dissolved carbon dioxide is very low in non-carbonate solutions. For example, green rust and magnetite were formed by the corrosion of cast iron at 50 °C in distilled, deionised water containing 0.4 p.p.m. of oxygen. Chemical analysis of the water gave the following results: before corrosion, $pH = 7.4$, alkalinity $= 5.0$ p.p.m.; after corrosion, $pH = 6.6$, alkalinity $= 5.5$ p.p.m.; alkalinity $= [CO_3{}^{2-}] + [HCO_3{}^-] + [CO_2] + [OH^-]$. Since, for H_2CO_3 at 50 °C, $K_1 = 5.19 \times 10^{-7}$ and $K_2 = 6.73 \times 10^{-11}$ (ref. 6) the bicarbonate ion is the predominant species in this *p*H range, and the carbonate ion concentration is extremely low.

In another experiment, green rust produced in sodium carbonate solution, $[CO_3{}^{2-}] = 200$ p.p.m. was examined using infrared spectroscopy. The presence of the hydroxyl group was clearly shown but the carbonate ion was not detected. In view of the comments of Brindley and Bish[5], a more extensive infrared spectroscopic investigation is being undertaken, but at present, our experiments do not support the view that the presence of the carbonate ion is particularly significant in the formation of green rusts by the corrosion of cast iron.

These and other points relating to processes involved in corrosion of cast iron will be developed elsewhere.

1 Mackay, A. L., *Nature*, **263**, 353 (1976).
2 Bernal, J. D., *Schweizer Arch. angew. Wiss. Tech.*, **26** 69 (1960).
3 McGill, I. R., McEnaney, B., and Smith, D. C., *Nature*, **259**, 200–201 (1976).
4 Pashley, D. W., *Epitaxial growth*, Part A (edit. by Matthews, J. W.), 2 (Academic, London, 1975).
5 Brindley, G. W., and Bish, D. L., *Nature*, **263**, 353 (1976).
6 Parsons, R., *Handbook of Electrochemical Constants*, 47 (Butterworths, London, 1959).

Production of heavy elements in neutron stars

CHECHETKIN and Kowalski[1] have proposed that the production of heavy elements in nature occurs by the ejection of matter from a neutron star surface. Their suggestion relies on the results of Chechetkin and Bisnovatyi-Kogan[2,3], who find such matter to be composed before ejection of 'jumbo' nuclei ($Z \gtrsim 160$; $A \gtrsim 640$). These results were derived using a highly uncertain extrapolation of binding energies into the extreme neutron-rich region and in the absence of fission. Their justification of the neglect of fission is based on a necessarily very crude estimate of fission lifetimes.

The nuclear stability against spontaneous fission is the result of a barrier in the potential energy curve of a nucleus as a function of deformation. In the framework of the macroscopic–microscopic method[4,5], the potential energy can be divided into two parts: (1) a macroscopic contribution that takes into account the smooth variation of the energy as a function of nucleon number and deformation, and (2) a microscopic contribution that includes shell effects which are significant only near the magic numbers for neutrons or protons. The macroscopic energy is calculated using the liquid-drop or droplet model of the nucleus, whereas the microscopic contribution is evaluated from the energy spectrum associated with an appropriate potential well.

As the proton number and the neutron excess of the nucleus are increased, the stability against fission resulting from macroscopic energy is reduced both by the increased repulsion of the Coulomb force with the addition of protons and by the decreased surface tension of the nucleus with the addition of neutrons. It is the rapid decrease of the macroscopic energy as a function of deformation that causes nuclei beyond the actinide region ($A \gtrsim 260$) to be unstable with respect to spontaneous fission. Beyond the actinide region, only strong single-particle effects near neutron or proton closed shells (as for the case of the superheavy island $A \simeq 298$, $N \simeq 184$, $Z \simeq 114$) can offer stability against spontaneous fission. In the immediate region surrounding the superheavy island, the fission barriers of nuclei are virtually non-existent[6].

Furthermore, in a neutron capture the neutron separation energy in the compound nucleus plus the kinetic energy of the incoming neutron easily lead to subsequent fission. Consequently, any process of successive neutron captures will be terminated by neutron-induced fission before significant spontaneous fission can occur. Recent calculations[6,7] of fission barrier heights and neutron separation energies in the *r*-process region indicate that the astrophysical *r* process

may well be terminated by neutron-induced fission in the mass region $A \sim 280$. It should, however, be pointed out that the exact point of this termination is very uncertain because of an extrapolation of surface-asymmetry and closed-shell effects from the region of known nuclei and because of the use of the Strutinsky procedure. The termination point may be somewhat higher ($A \sim 350$), consistent with the possible discovery of superheavy elements in pleochroic halo inclusions[8] and their production in a conventional *r* process[9]. In spite of the uncertainties of the calculations of fission barriers[6,7], we believe that the existence of nuclei as large as $A = 640$, $N = 480$, $Z = 160$ is definitely ruled out. We should also mention that similarly large nuclei had been predicted in equilibrium conditions in neutron star matter[10,11]. This, however, has also been conclusively shown to be incorrect[12–14]. In conclusion, the nuclear basis of Chechetkin and Kowalski's suggestion for the production of superheavy elements is unfounded although it may well be that these elements can be produced in an *r* process occurring in the surface material of neutron stars. Studies along these lines are under way.

J.- ROBERT BUCHLER

Department of Physics and Astronomy,
University of Florida,
Gainesville, Florida 32611

WILLIAM A. FOWLER
MICHAEL J. NEWMAN

W. K. Kellogg Radiation Laboratory,
California Institute of Technology,
Pasadena, California 91125

MICHAEL HOWARD

Department of Physics,
Northwestern University,
Evanston, Illinois 60201

1 Chechetkin, V. M., and Kowalski, M., *Nature*, **249**, 643–644 (1976).
2 Chechetkin, V. M., *Soviet Astr.*, **15**, 45–50 (1971).
3 Bisnovatyi-Kogan, G. S., and Chechetkin, V. M., *Astrophys. Space Sci.*, **26**, 25–46 (1974).
4 Nix, J. R., *A. Rev. nucl. Sci.*, **22**, 65–120 (1972).
5 Möller, P., and Nix, J. R., *Proc. Third IAEA Symp. Physics and Chemistry of Fission, Rochester, 1973*, I, 103 (IAEA, Vienna, 1974).
6 Howard, W. M., and Nix, J. R., *ibid.*, 145.
7 Bengtsson, R., Boleu, R., Larsson, S. E., and Randrup, J., *Proc. Nobel Symposium on Superheavy Elements, Ronneby, Sweden, 1973*, 10A (1974). (edit. by Nilsson S. G., and Nilsson, N. R.) (Alunquist and Wiksell, 1974).
8 Gentry, R. V., *et al.*, *Phys. Rev. Lett.*, **37**, 11 (1976).
9 Schramm, D. N., and Fowler, W. A., *Nature*, **231**, 103–106 (1971).
10 Bethe, H. A., Börner, G., and Sato, K., *Astr. Astrophys.*, **7**, 279–288 (1970).
11 Baym, G., Bethe, H. A., and Pethick, C. J., *Nucl. Phys.*, A175, 225–271 (1971).
12 Buchler, J.-R., and Barkat, Z. K., *Phys. Rev. Lett.*, **27**, 48–51 (1971).
13 Barkat, Z. K., Buchler, J.-R., and Ingber, L., *Astrophys. J.*, **176**, 723–738 (1972).
14 Negele, J. W., and Vautherin, D., *Nucl. Phys.*, A207, 298–320 (1973).

reviews

Assuming the mantle of soothsayer

L. Knopoff

Earthquake Prediction. (Developments in Solid Earth Geophysics, Vol. 9.) By Tsuneji Rikitake. Pp. xvi+357. (Elsevier Scientific: Amsterdam, Oxford and New York, 1976.) Dfl.95; $37.95.

STRONG earthquakes hold such terror for unprotected populations that those capable of forecasting earthquakes with any degree of reliability whatsoever, have a ready and appreciative audience. Thus, earthquake prediction has been the business of psychics and astrologers for centuries. The unhappy consequences of recent severe earthquakes have inspired significant effort in four nations to spur science to assume the mantle of soothsayer. Government-sponsored prediction research has been undertaken on a major scale in the USSR (since 1949), Japan (1963), China (1966) and the US (1971). In these countries, concerted efforts have been made in what must be the first priority in any programme—obtaining base-line information. These activities have been carried out to varying degrees in each country by measurement of strains, geodetic measurements, observations of micro-earthquake activity, geomagnetic, geo-electric and elastic properties, and collection of geochemical information.

Each of the four programs has its own emphasis. Professor Rikitake has been a leader of the Japanese program, and writes as an authority eminently qualified to review the Japanese program. The longest chapter in the book (62 pages) is concerned with geodetic surveying methods for observing precursory deformation of the Earth's surface and summaries of observations of such deformation before a number of earthquake events. Although geodetic observations are comparatively lightly emphasised in the other programs, they are a cornerstone of the Japanese national program.

The reader should not expect to receive a broad exposure to the basic problems of earthquake prediction or to the arguments and uncertainties that pervade the subject. Instead, this book provides a thorough review of the phenomenology of observations deemed pertinent to prediction. The reader who expects to derive understanding of the processes that underly the observations should look elsewhere.

The physical basis of earthquake prediction is unsatisfactorily treated. The shortest chapter in the book (4 pages) is concerned with the dilatancy model of the focal zone; in this model, the volume of matter increases due to the formation of small cracks shortly before rupture. Dilatancy is the only physical model proposed so far that has the promise of providing an understanding for the various precursory phenomena. Professor Rikitake withdraws from a detailed discussion of dilatancy, including the controversy over the role of water as an active agent in the dilatancy process.

The lengthy chapter entitled "The Theory of Earthquake Prediction", has no theory therein, but is concerned instead with statistical approaches to earthquake prediction plus an extensive catalogue of the events for which precursors have been observed. Two or three pages of the book are devoted to the phenomenology of brittle fracture and strike-slip processes; nothing is presented describing the physical basis in fracture mechanics for the earthquake event.

The book is well written in language understandable to the non-expert. The non-idiomatic use of the English language is not offensive. The editing and the selection of topics make the book a highly personal view of earthquake prediction. Major aspects of the topic deserve to be presented more broadly than is done here. In spite of extraordinary omissions, this book is useful for the extensive tabulations of observations of diverse precursory phenomena. ☐

L. Knopoff is Professor of Geophysics at the Institute of Geophysics and Planetary Physics, University of California, Los Angeles, California.

Magnetic resonance techniques

Magnetic Resonance of Biomolecules: An Introduction to the Theory and Practice of NMR and ESR in Biological Systems. By P. F. Knowles, D. Marsh and H. W. E. Rattle. Pp. 343. (Wiley-Interscience: New York and London, May, 1976.) Cloth £9.75, $19.75; paper £4.25, $8.75.

THE applications of magnetic resonance techniques to biological problems are now of sufficient importance to merit inclusion in any well-balanced undergraduate degree course in biochemistry or biophysics. Although there are several specialist books dealing with such applications there has been no text written primarily for the undergraduate student.

Drs Knowles, Marsh and Rattle have now produced such a book dealing with the applications of nuclear magnetic resonance (NMR) and electron spin resonance (ESR) spectroscopy to biological problems. On the whole, they have produced a very useful book for both student and teacher alike. The parts of the book dealing with ESR are particularly valuable because such teaching material is not easily accessible elsewhere.

The basic theory of NMR and ESR are discussed together in an opening chapter and other chapters deal with the spectral parameters, practical aspects and biological applications of each technique. The carefully selected material is presented in an interesting manner and due emphasis is placed on the advanced experimental techniques which are required for successful biological studies. In general the theoretical aspects have been dealt with adequately although I would have preferred to see the theory of NMR and ESR introduced separately. In the NMR section a more detailed account of chemical exchange and relaxation behaviour would have been useful in view of the importance of these aspects of the theory in biological studies.

In some of the examples it was not always clear how the biological information had been derived from the NMR data. For example, a conformational structure for oxytocin was presented but none of the NMR argu-

ments leading to this were given. Obviously in a text of this type the authors have had to be very selective in their choice of examples of the numerous applications; they have succeeded admirably in finding examples which illustrate the scope of the techniques as well as providing interesting biological information.

The text would have been improved by including some consideration of conformational studies of small molecules of biological interest using spin–spin coupling constants and some studies of ligand binding to proteins. In general however the authors have provided an impressive coverage of the major areas of application including studies of protein and transfer RNA conformations and structures, protein unfolding, protein–protein interactions, protein hydration, metalloproteins, spin-labelled proteins and studies of phospholipid membranes.

This book can certainly be recommended to students and teachers wishing to learn more about the usefulness of magnetic resonance techniques in biology.   **J. Feeney**

Dr Feeney is a member of the scientific staff in the Division of Molecular Pharmacology at the National Institute of Medical Research, London, UK.

Immunology of reproduction

Immunobiology of Trobophlast. (Clinical and Experimental Immunoreproduction, 1.) Edited by R. G. Edwards, C. W. S. Howe and M. H. Johnson. Pp. x+284. (Cambridge University: Cambridge, London and New York, August 1975.) £6.

THIS volume represents the papers and discussion of a meeting held in January 1974 in the Physiology Laboratory, Cambridge, UK, and is the first of a series of meetings held on specific topics of immunoreproduction in association with the activities of the International Coordination Committee for the Immunology of Reproduction. All the participants were either British or were visiting Britain.

The papers all, in one way or another, address themselves to the key fact that the foetus is able to survive in the uterus in spite of the antigenic disparity which by rights should lead to its rejection as an allograft. Papers by A. C. Allison, by M. H. Johnson, by K. D. Bagshaw and S. Lawler, and by W. P. Faulk *et al.* document the complex antigenicity of the developing trophoblast, placenta and foetus. Details of trophoblast differentiation are described by R. L. Gardner and by W. D. Billington. The trophoblast and placental function in prevention of cell traffic and transfer of antigens and antibody is discussed by M. Adinolfi, whereas endocrine effects on the immunological consequences of implantation of the antigenically foreign foetus are described by R. Borland *et al.* W. R. Allen reports on the complex endocrinological and immunological features accompanying implantation and foetal development in the mare. Two extensive papers, one an overview of lymphocyte physiology by C. W. S. Howe, the other a review of the complex and often divergent observation on the effect of antigenic disparity on placental size, implantation and embryonic survival of mammalian embryos by A. McLaren complete the volume.

It is remarkable that in spite of the limited number of papers, the volume represents well the current state of knowledge in this field. The more comprehensive papers are doubtlessly the most valuable in that they focus in depth on the problems of trophoblast immunobiology, and in depth consideration of such problems is a luxury permitted all too infrequently in a time of increasing page charges and decreasing page allowances. The additional feature of publishing the discussion comments following each paper is especially welcome, for it provides the reader with a feeling for the current state of acceptance of or disagreement with the views expressed by the authors.

The volume is a most valuable addition to the literature in the fields of reproductive biology and of immunology.   **Robert Auerbach**

Robert Auerbach is Professor of Zoology at the University of Wisconsin, Madison, Wisconsin.

Membrane series

Biological Membranes. Vol. 3. Edited by Dennis Chapman and Donald F. H. Wallach. Pp. x+362. (Academic: London and New York, May 1976.) £11; $27.75.

WHEN *Biological Membranes, Physical Fact and Function* was published in 1968 there was no indication that it was to become volume 1 of a series which has now reached volume 3 and seems likely to continue. The publication of such volumes at irregular intervals of several years tends to delay some of the contributions unfairly and to provide an irregular and unpredictable coverage of the field. Nevertheless, in the book under review, the editors have collected together a group of six interesting topics and appropriate authors, most of whom seem to have written their contributions during 1974. The chapters are, in general, authoritative and well presented, and on such a variety of topics as to provide something of interest for everyone. Most libraries and many individuals may therefore be persuaded to continue to collect the series.

Two of the new topics, concerned with neutron diffraction and scattering and with the use of lanthanides to probe the role of calcium in membranes, are still at the exploratory stage and are reviewed largely to emphasise their potential for membrane studies. Other chapters discuss lipid and protein mobilities in membranes and the exchange of lipids between membranes and lipoproteins. The two remaining chapters provide a comprehensive review of nuclear membrane structure and function, and a specialist assessment of the specificity and significance of interactions between plasma membranes and gammaglobulins. The author index is four times as long as the subject index even though the latter includes all but three of the rare earth elements (all listed on p154).   **J. B. Finean**

Dr Finean is a Reader in Molecular Biology in the Department of Biochemistry at the University of Birmingham, UK.

Light scattering

Dynamic Light Scattering. By Bruce J. Berne and Robert Pecora. (Wiley: New York, March 1976.) £14.00; $27.70.

CLOSELY following the introduction of the laser we have in recent years witnessed a considerable increase in the number of original publications and review articles on light scattering theory and its applications to chemistry, biology and physics. The novel feature which transformed classical light scattering into a modern field of study was the possibility of undertaking dynamic studies over an extensive range of time and frequency scales. Not all expectations have been fulfiled; the study of chemically reacting systems for instance has been rather disappointing so far and the successful extraction of higher order information, beyond the derivation of translational diffusion constants in dilute macromolecular solutions, has been the rare exception rather than the commonplace event.

From the book under review, as well as from its recent predecessor (see *Nature*, **255**, 90, 1975) we learn that sufficient time has not yet elapsed to place the whole field into proper perspective. This reviewer would have been grateful for a logical development of modern scattering theory comprising the phenomenological, electromagnetic and molecular aspects. The authors are eminently qualified for this undertaking. They present a welcome wealth of theoretical derivations, sometimes on a very elementary and at other times on a rather advanced specialised level. Interspersion of a large number of applications from many different fields detracts from the general purpose. Whereas the theory could easily have been presented as a permanent whole, the applications are fragmentary and ephemeral, and can only be fully appreciated by reference to the original works. The thread is broken and the consistency lost.

There are a number of details which I believe can be easily corrected in a future reprint. Some terms and names are misspelt, some references in the text are incomplete, and some references are given as 1974 and 1975 as being in press—in a book published in 1976. A glossary of symbols would have been helpful in view of the complicated nomenclature. To find out which diffusion coefficient is determined in the light scattering experiment the reader must struggle through the whole development of non-equilibrium thermodynamics taken from another source.

In spite of the above reservations, this book will be useful as a sourcebook of important theoretical information in specialised research laboratories active in a wide range of scattering studies. **Henryk Eisenberg**

Henryk Eisenberg is a member of the Polymer Department, The Weizmann Institute of Science, Rehovet, Israel.

Pion–pion interaction

Pion-Pion Interactions in Particle Physics. By B. R. Martin, D. Morgan and G. Shaw. Pp. xi+460. (Academic: London and New York, May 1976.) £16; $40.50.

AT some time during their career, most high energy physicists must obtain some acquaintance with the pion-pion interaction, because of its pervasive nature in hadronic physics. In future, I will have no hesitation in recommending the appropriate chapter of this book to them as an essential reference. The book provides a timely review of the subject, as there have been many important developments, both experimental and theoretical, since the appearance of Petersen's comprehensive *Physics Reports* article in 1971.

The lack of real pion targets means that special techniques have been developed to extract data on $\pi\pi$ scattering. A substantial part of this book is devoted to a careful explanation of these methods and the results from recent high precision experiments. The other major section of the book deals with theories and models of $\pi\pi$ scattering.

In places the book inevitably reflects the authors' own approach to the field. It is, for example, somewhat surprising that explicit expressions for the S and P wave contributions to the important Roy equations are not included in the otherwise extensive appendices. In general, however, they are to be congratulated on the care they have taken in preparing this book, which is destined to become a standard reference and should be available in all high energy physics libraries. **C. D. Froggatt**

Dr C. D. Froggatt is a Lecturer in the Department of Natural Philosophy, Glasgow University and has made many contributions to the phenomenology of pion–pion scattering.

obituary

Nikolai Ivanovich Muskhelishvili the eminent mathematician died on July 16. He was born on February 16, 1891 into a remarkable family, of great renown in Georgian history. His father, Ivan Levanovich, was a General in the Engineering Corps of the Imperial Russian Army, and had interesting ideas on the education of children. He spent a good deal of time teaching his own children—particularly 'Niko' whose mathematical ability he recognised at an early age. Muskhelishvili's mother, too, was a person of great culture as well as charm.

After completing his secondary education at the Second Tbilisi High School, Muskhelishvili entered the Faculty of Physics and Mathematics of the University of St. Petersburg in 1909. His outstanding ability early attracted attention, and legends were soon circulating among the student body of the accomplishments of the remarkably gifted young Georgian. He graduated in 1914 and in the following year presented a diploma thesis of such merit that he was invited to become research assistant to Gurii Vasil'evich Kolosov, a pioneer in the application of function theory to plane problems in elasticity. It was Kolosov who guided Muskhelishvili's first steps in the subject which in the next sixty years he was to revolutionise and dominate, and, in the process, to influence critically much of the development of mathematics (both pure and applied) in the USSR. One of his first papers, *Sur l'intégration de l'équation biharmonique,* published in 1919, gave a foretaste of the simplicity and elegance which were soon to become the distinguishing features of everything he wrote. As with most of his later writings on elasticity, the problem considered was of great practical as well as mathematical interest, for a special case considered by Muskhelishvili proved to be basic to the theory of brittle fracture in thin plates.

In 1917 Muskhelishvili was appointed an assistant in what was now the University of Petrograd and two years later was promoted to an instructorship. It is typical of him that in the years 1915–1920 he gave a great deal of his time and energy to teaching in other institutions of higher learning in Petrograd which, because of the troubled times, were experiencing grave staffing problems.

When the university of Tbilisi was founded, he was invited to join the faculty. He returned to his homeland in 1920 and remained there until his death. Soon after his return to Georgia, Muskhelishvili assumed the role of leader of Georgian mathematics and created in Tbilisi one of the chief mathematical centres in the world. Initially, the group concentrated on the mathematical theory of elasticity, but later, comparable effort was put into important fields of analysis. Many of the famous names of Soviet mathematics appear in the records of the Tbilisi group and many more were influenced by the great stream of papers and research monographs which flowed from it.

Muskhelishvili's own research contributions were almost entirely in the theory of plane elasticity and in the theory of singular integral equations, but it was not only those branches of mathematics most directly connected with his own research interests that he sponsored. To take only one example: it was through his active support that the Georgian school of topology was founded and nurtured in its early years.

During this period of intense scientific activity extending over half a century, Muskhelishvili was tireless in his efforts to organise efficiently scientific activities in Georgia. He gave devoted service to the Academy of Sciences of the Georgian SSR, and from his election in 1939 as a full member of the Academy of Sciences of the USSR he served repeatedly on its Presidium. When the National Committee of the USSR for Theoretical and Applied Mechanics was established in 1957 he was the obvious choice for its Chairman and he served long in that capacity, with the representation of the USSR on various international bodies which that office involves.

Muskhelishvili was influential in wider circles than the scientific. Believing that a university professor should participate fully in the life of his country he was active in the political field; among the offices he held was Deputy of the Supreme Soviet of the USSR.

For his scientific work and for the significant role he has played in the development of Soviet science, Muskhelishvili received many awards and honours, from his own government and from foreign institutions. Those who were privileged to know him and his family values the memory of his friendship and his gaiety, recall with affection the delights of his companionship as much as his incisive judgment on scientific questions. Future generations will know him only from his research papers and his two great monographs *Basic Problems in the Mathematical Theory of Elasticity* and *Singular Integral Equations* which achieve the seemingly incompatible aims of creating a satisfying unified theory and at the same time pointing the way to new realms of enquiry.

In all his published work Muskhelishvili showed that he was both a pure and an applied mathematician. The 'applied' aspect of his nature was shown by his willingness to look at problems of real practical significance, the 'pure' in the rigorous analysis to which he subjected the problems thrown up by his investigations into mechanics.

I. N. Sneddon

Take your time, Mr Benn

LIKE a good thriller, the Royal Commission on Environmental Pollution's report on the environmental risks of nuclear power, published last week, gives relatively little away in its early chapters. With the greatest of care (and surely such a readable report deserves a wider audience than the Stationery Office is accustomed to delivering) the scene is set in what could serve as an intelligent layman's guide to nuclear energy. A little pat on the back here, a modest rebuke there; a new committee suggested, a possible gap in interdepartmental responsibilities pointed out. Then, with the end of the book almost in sight the action quickens—in a somewhat unpredictable way—and the final pages are highly controversial.

The commission, chaired for the duration of the preparation of this report by Sir Brian Flowers (though now headed by Professor Hans Kornberg), has as its remit to advise on environmental pollution, on research, and on future dangers to the environment. The projected growth of the civil nuclear power programme, with all the associated general concern over possible radiological hazards, stands out as an appropriate field of study for the commission, particularly with pressures growing for an early government decision on major investment in a demonstration commercial fast-breeder reactor (CFR). The report has come none to soon; authorisation to go ahead with the design for the first CFR—popularly known as CFR-1—was being sought from the Energy Secretary, Mr Benn, by the United Kingdom Atomic Energy Authority (UKAEA) within a matter of weeks. At least the publication of the report will give Mr Benn good cause to delay his decision. Careful analysis of the commission's arguments could conceivably persuade him that clearance for CFR-1 would be an unwise option. For the widespread implications of such commitment, even to just one CFR, will greatly overshadow even the implications, little realised at the time, of commitment to building Concorde.

At first glance the benefits of fast breeder reactors seem overwhelmingly attractive. All present-day thermal nuclear reactors are fuelled primarily by fissile uranium-235. Even so, non-fissile uranium-238 is by far the dominant isotope in the fuel elements, and thermal reactors convert some of this to plutonium-239 which, like uranium-235, is fissile. But only a small amount is converted. This is where fast breeder reactors steal an advantage: they "breed" more fissile material in the form of plutonium-239 than is lost in the form of uranium-235. Obviously, this requires a major rethink of reactor design. A high density of fast, rather than slow, "thermal", neutrons is required, and the reactor core must be very compact. Consequently, the starting fuel must be a mixture of plutonium and uranium. The engineering difficulties are not insuperable; already prototypes have been built or are under construction, not only in the UK, but also in France, the Soviet Union, the US, and West Germany. Full-scale, commercial fast breeders would probably run at above the 1,000 MW mark, would be cooled by liquid sodium, and would, after the initial charge of fissile uranium and plutonium, never need specially enriched fuel, being able not only to run on depleted fuel from thermal reactors but also to feed back fissile plutonium to thermal reactors. All of which looks immensely attractive as a medium term prospect for nuclear power stations.

Except for one catch : all the plutonium which will be around. By the early 21st century, according to some scenarios, the world will be geared to a plutonium economy, with the developed nations satisfying their energy requirements by using and circulating plutonium—a highly toxic metal quite easily convertible into nuclear weaponry. The commission accordingly devotes much of its attention to the problems that the plutonium economy throws up. It also, however, uses the opportunity of the report to review many other matters, such as radiological standards and waste management, which still needs serious attention regardless of any decision on the CFR.

The report first reviews what is known in radiobiology, with particular attention to plutonium. The present maximum permissible body burden is 0.6 mg, and the commission does not consider this standard to be seriously in error. In reaching this decision it had to consider the "hot particle theory", Gofman's analysis of the effects of smoking on lung clearance rates, and assertions that deaths from leukaemia among workers at the Windscale reprocessing plant in Cumbria were significantly high. For the first two, independent consultants were brought in, for the third the National Radiological Protection Board (NRPB) produced a paper, and, as a result, fears of any serious error in radiological standards were dismissed in every case. The consultants' reports have not been published nor even promised, and many will want to suspend judgment until they are openly available.

Next, the commission looks at how radiation standards are arrived at and enforced. The way that the International Commission on Radiological Protection reaches decisions is deemed satisfactory—the UK already fully accepts these standards. But there is some feeling that the NRPB, "a national point of authoritative reference in radiological protection", is not fulfilling a sufficiently central role, largely because it is fairly new while other bodies concerned with nuclear

power still have much in-house expertise. The commission recommends that consultation with the NRPB should become mandatory and that the NRPB should be statutorily responsible for endorsing basic standards. Further, the NRPB should provide the focal point for coordinating research into radioactivity in the environment, which is at present well done in the marine environment by the Fisheries Research Laboratory, but only patchily covered in the atmospheric and terrestrial environments. More work is also pointed towards the NRPB in the monitoring of radiation workers and of discharges.

If the comments on standards are relatively soothing, the pace of the report quickens a little when it comes to waste management. It is here, of course, that emotive issues—"leaving a radioactive time-bomb as a legacy for countless generations" and so on—could crop up. The commission asserts that "there should be no commitment to a large programme of nuclear fission power until it has been demonstrated beyond reasonable doubt that a method exists to ensure the safe containment of long-lived, highly radioactive waste for the indefinite future". It also makes it clear that the method does not yet exist. Lower levels of radioactivity are also considered and one or two comments made, particularly about the less than highest standards of housekeeping observed during the commission's 1974 visit to Windscale. But it is the high level wastes, mainly fission products, at present in tanks at Windscale which get most attention. Vitrification is seen as indispensible; development started in the late 1950s, but will not be commercially available until 1985 (there was a strange period of inactivity in the 1960s). Final disposal of the 15-tonne cylinders (a magnox reactor generates three tonnes a year in the UK) is the real problem.

The commission quickly rule out space rockets, subduction at ocean trenches, deep burial in Antarctica, and the time-honoured means of disposing of ships' waste—over the side. Burial is seen as essential, either on land or down specially drilled holes at sea. Surprisingly little research has been done in the UK on these options (maybe tHat is why they still seem attractive!), and the commission is critical both of UKAEA and British Nuclear Fuels Limited for their complacency. It proposes that a Nuclear Waste Disposal Corporation be formed to get on with the job.

The two issues that get most headlines in connection with plans for widespread nuclear development are the possibility of serious reactor accidents and the risk of illicit activities. In both of these fields the commission acknowledges itself to be straying well outside its brief, so its views can only really be those of a collection of very well informed laymen. Nonetheless, the chapters on these subjects, particularly on the latter, are important in that they provide a more total approach than is likely to appear in official discussions elsewhere.

Any assessment of the threat that terrorists could pose if a plutonium economy becomes established, or that security measures could pose to civil liberties is extremely difficult, and the commission rightly does little more than spell out some of the problems— it provides no cut and dried answers. On balance, the likely effects on civil liberties do not on the surface look serious in a generation used to all sorts of screening. But Sir Brian Flowers has already remarked that every wedge has its thin end, and even while the present arming of nuclear security guards must be seen as an eminently sensible move, the ultimate long-term implications need careful assessment. The terrorist threat looks more obviously bleaker. Would it be possible to site a fast-breeder in Northern Ireland? Is the rest of the UK immune from terrorist threats, bearing in mind that terrorist groups could be fed information by other states with nuclear capabilities? And the question is inevitably a world problem—if the UK opts for a plutonium economy it may significantly encourage other countries with less political stability to do likewise; even a terrorist attack in just one country would have worldwide repercussions. Still, the report carefully avoids judging these complex human questions, merely listing the potential dangers. But then, in a quite unexpected development, it proceeds to cast severe doubts on fast breeders.

The sudden switch comes in the chapter on energy strategy, and the argument runs like this. The Department of Energy (DEN) has made various energy demand forecasts for the next 50 years based on reasonable growth assumptions. In the most favoured current version electricity consumption more than trebles by 2025; sumption more than trebles by 2025: 104 GW of nuclear electricity are needed by 2000 rising to 365 GW 25 years later. The commission observes that emphasis on electricity ensures that
● There will be a huge call for investment in the energy sector (3.4% of GNP, as opposed to 1.4% today),
● A large number of coastal sites must be developed by 2030—a requirement that it may not be possible to meet,
● A lot of the primary energy converted will be wasted in the form of low grade heat.

In a couple of breathtaking pages the commission becomes merchant banker and declares itself "very doubtful" on the investment prospects, becomes the Fine Arts Commission and declares all these power stations a substantial adverse effect on amenity, and reverts to environmental commission to declare in three paragraphs that the heat might affect the climate. There thus appear "very considerable environmental objections to the high-nuclear, high-electric future". The feeling is rather as if a referee, after arbitrating over a goalless draw for 89 minutes, suddenly decides to boot the ball into the net to break the deadlock. Once the label "environmental objections" has been pinned on to one DEN scenario (although many other scenarios have been published—even down to a non-nuclear future) the rest of the report becomes coloured by a general optimism on alternative sources. Coal, of course, would bear the brunt of the load; and it is noticeable how the immense radiological protection lavished on nuclear workers is not matched by any comparable health concern for the miners who will have to dig the coal up.

One feels that with one hasty glance at 2025 the commission has made its mind up on the delicate issue of fast-breeder reactors—they should be introduced "only if demonstrably essential"; in the meantime money should be directed towards other technologies, notably fusion (though it is not clear that fusion research would benefit from a diversion of even a little of the £2,000 million that might eventually go into CFR-1), and international collaboration should be sought out for fast breeder development.

So where does this leave Mr Benn? The nuclear industry claims that it needs early decisions; the commission urges delay. Even though some of the commission's otherwise excellent reasoning is marred by sloppy thinking, the issues raised must give Mr Benn pause for thought. The question that undoubtedly is central in his mind is whether the technological momentum that will be built up by commitment to CFR-1 is stoppable if the warning bells begin to ring. Unless he can devise ways of ensuring this, and until the nuclear industry can demonstrate its understanding of the need for aborting mechanisms, he would do well to adopt a delaying strategy. The middle course steered by the commission makes one thing outstandingly clear: there is need for more discussion and a greater appraisal of the possibilities. Is an early decision on CFR-1 really as essential as the reactor itself may ultimately prove? The extra couple of months already gained since the report's publication (the DEN last week announced that the decision planned for October will now wait until at least the end of the year) must be seen to be more than just a token delay. □

Soviet dissidents (1)

He who would dissident be

Why do so many Soviet scientists become dissidents
and refusniks? **Vera Rich** considers the question

ONE of the paradoxes of contemporary Soviet 'opposition', whether it comes from dissidents proper, who campaign for improvement of the system from within, or from the Jewish refusniks, whose only wish regarding the Soviet Union is to leave it for Israel, is the presence of large numbers of those who might be expected to benefit most from the *status quo*—scientists. In a culture which bases its whole economic policy on the "implementation of the scientific technological revolution" and which, with its total state control of all resources, can allot virtually unlimited funds to a favoured research project, the position of a scientist, one might feel, is surely such as to ensure all the fringe benefits which Soviet society affords its elite—access to privileged shops, better housing, Black Sea holidays, and the like. Yet, by the kind of paradox which characterises so much of Russian history, it is partly this privileged position that has led to the development of dissidence among scientists.

The story begins back in the late 1940s when Stalin, having 'abolished' Mendelian genetics throughout the Soviet Union, and replaced it with the peculiar notions associated with the name of Lysenko, turned his attention to modern physics. Since the basic tenets of relativity theory and quantum mechanics seemed, in the opinion of certain favoured theoreticians, contrary to Marxism–Leninism, Stalin aimed to confine physics research to what was congruous with Newtonian mechanics. But a small group of physicists, notably Kurchatov, Kapitsa and Vavilov, managed to convince him that without modern physics he could expect neither nuclear power, nor, and this was the clinching argument, a nuclear arsenal. Modern physics was accordingly permitted to continue, although, until Stalin's death, textbooks were liable to contain face-saving clauses to the effect, for example, that the second law of thermodynamics was "a local phenomenon in this part of the universe".

Having protected their own discipline from destruction, it was perhaps natural that physicists should come to the aid of their less-fortunate scientific brethren, the harassed geneticists, in their fight to re-establish genetics as a valid field of research. Lev Tumerman, a physicist working on the luminescence of organic molecules, was instrumental in arranging the first

seminar on chromosomes, held in the Lebedev Physics Institute during the late 1950s, while in 1959, Timofeev-Resovskii gave his first genetics lecture in Kapitsa's Institute. And in the great debate of 1964, which finally crushed the attempted resurgence of Lysenkoism and re-established genetics as a branch of Soviet science (although it did not receive 'priority' funding for another 10 years), certain physicists, notably Andrei Sakharov, played a significant role.

From the defence of a single scientific discipline to a deep involvement with human rights as a whole is, however, a great step to take. But a number of factors in Soviet society established a climate for dissent. The phenomenon of *samizdat*—'do-it-yourself' publishing and distribution, to sidestep official censorship—reaches far back into the 19th and 18th centuries, and it received a new impetus with the mushrooming availability of typewriters and, on occasion, photocopiers. The education reforms of 1958 which, in effect, made Russian virtually the only language of higher education and academic life, produced great discontent among the non-Russian republics of the Union, notably Ukraine and the Baltic States. The highly anti-Israeli attitude of the Soviet media at the time of the Six-Day War for the first time impressed upon Jewish intellectuals an awareness that Israel was not merely a paper creation (on the lines of the Soviet 'Jewish homeland' of Biribidjan). Israel at last came to be seen as a viable alternative society where Soviet Jews might live and work unhampered by the many disabilities which (*de facto*, and in contradiction to the Soviet constitution) are still the lot of those with "Jewish nationality" stamped in their internal Soviet passport.

But *samizdat* was, originally, the province only of the more daring creative writers, while the problems of minority nationalities had always been endemic to the multinational Russian empire. And except in a few cases, such as the promising Ukrainian school of cybernetics, little was lost to Soviet research when the language of instruction and publication was officially changed—a scientist, after all, must expect to read and, on occasion, to publish, papers in languages other than his own. Why then this great participation by the scientists in the human rights movement and dissidence?

Kronid Lyubarsky, 5-year sentence

Andrei Sakharov, the acknowledged leader of scientific dissidence, has given his own account. Though he originally viewed nuclear balance as an important factor of world peace, he subsequently began, as early as 1958, to agitate for the cessation of nuclear testing. The reason for the switch was clear. He became convinced that no further benefits could be gained from such tests, believing that the sole outcome would be nothing more than an increase in the natural level of genetic hazard. The fundamental change of attitude undergone by Sakharov by the early 1960s led to his concern about the 'criminal nature' not only of the tests but of the whole concept of nuclear weaponry. Beset with a sense of helplessness he devoted an ever-increasing amount of time first to *samizdat* writings (the 1968 monograph *Thoughts on Progress, Coexistence, and Intellectual Freedom*, and the *Second Manifesto* of 1970), and later to appeals on behalf of fellow dissidents threatened or imprisoned for their views and activities.

There is no doubt that Sakharov's personal charisma became a major factor in drawing together the little group who formed his illicit 'human rights' movement, and its subsequent off-shoots—the Moscow branch of Amnesty International and the 'Helsinki monitoring' group. Among Sakharov's associates are such leading lights as the physicists Valentin Turchin and Andrei Tverdokhlebov, the mathematician Valerii Chalidze, and the biologist Sergei Kovalev.

But what of the many others who did not at first have any personal contact with the great names of the dissident world? Some general trends may be distinguished. First, many students turn originally to science not so much from a personal preference for the subject as such, nor from hopes of a prestige job, but because science offers the hope of academic freedom, so

patently absent from the humanities which are dominated by the hard-and-fast requirements of Marxist–Leninist ideology. This expectation may remain unfulfilled—Soviet medical students are often disillusioned with psychiatry and psychology lectures based on Pavlovian behaviourism and a rigid definition of what constitutes 'normality'. Moreover, even in the absence of such specific disillusionment, the Soviet scientific establishment itself produces an atmosphere of frustration.

Travel to conferences abroad (or even, in some cases, within the Soviet Union) is frequently blocked by barriers of officialdom open only to the favoured few. Foreign publications cannot be obtained on subscription and are often inaccessible in libraries except to those with special security clearance.

The classic story of frustration recounted in *The Medvedev Papers* is by no means unique. The Soviet system makes no allowance for serendipity—a frequent complaint from young researchers is : "They expected me to order all the equipment and reagents before I started, but if I'd known exactly what I wanted, I wouldn't have needed to do the experiment"; or, "The official supply system is useless. You can only get on if your laboratory has a good 'fixer' who can arrange a swap with someone who has what you need." (This latter difficulty leads to deliberate over-ordering and the development of a system of polygonal bartering, involving several laboratories at a time.)

Such frustrations may eventually lead scientists to work more actively for human rights, even at the cost of a secure career. Vladimir Bukovskii, a biology student expelled from Moscow University for his dissident views, became the first prominent campaigner to focus world attention on the misuse of psychiatry by a regime aiming to equate dissidence with insanity. At the very least the frustration which so motivated Bukovskii may lead to a state of mind describable as "passive dissidence", creating a ready readership for available *samizdat* material. It should be remembered, of course, that even the temporary possession of *samizdat* material is a serious offence. Biologist Nina Strokata-Karavan'ska, for example, served a four year sentence for possession, and is still exiled from her native Ukraine. Similarly, Sergei Kovalev, another biologist, is serving a 7-year sentence in a strict regime camp. (Kovalev, a member of the Moscow Amnesty group was curiously charged with possession of the illegal *Chronicle of the Lithuanian Catholic Church*, although he is neither a Lithuanian, nor, so far as is known, a Catholic). Even greater penalties can be involved. Astrophysicist Kronid Lyubarskii, is not only serving a five year sentence—

moves are now afoot to deprive him of his academic degree.

For a Jewish scientist, however, discontent typically results in an application to emigrate to Israel. (As a convenient one-way ticket out of the Soviet Union, this opportunity is occasionally afforded to troublesome non-Jews, such as Leonid Plyushch, the Ukrainian mathematician.) But emigration is by no means an easy option. Successful applicants often indicate that the obstacles placed in their way actually intensify their Jewish consciousness, so that Israel becomes a positive permanent goal rather than merely the only available sanctuary. Once an application is placed, a whole chain of sanctions are imposed. Job dismissal is virtually automatic—with a consequent long-term loss of access to all professional publications and loss of contact with new specialist developments. (It was in an attempt to maintain at least some semblance of scientific life that Aleksandr Voronel founded the famous Sunday seminars). The authorities can throw up a whole battery of reasons for refusing exit visas. These include military service (often invoked in the case of over-age invalids, such as Mark Azbel, Voronel's successor as leader of the Sunday seminars); security restrictions (even in fields where the Soviet Union is not noticeably ahead of, or even abreast with, developments elsewhere): and the notorious 'education tax' (waived since the Nixon visit to Moscow, but never actually revoked) which required the payment of an extortionate sum, allegedly equal to the estimated cost of an applicant's university and post-graduate studies.

Reprisals may extend to relatives. Metallurgist Evgenii Reinberg was expelled from the Communist Party, and subsequently from his academic post, when his son applied for a visa. Endocrinologist Mikhail Shtern, now serving an 8-year sentence for "accepting bribes", was actually told that the accusation was made "in connection with your family's desire to emigrate". And, of course, dismissal from one's job involves the constant threat of arrest for 'parasitism', or being without visible means of support, a threat directed particularly often, it appears, at the members of Azbel's seminar.

Needless to say, a refusnik scientist shares the same disabilities as a dissident where international conferences are concerned. The recent absence of Benor Gurfel from the European meeting of the Econometrics Society in Helsinki, and of Academician Veniamin Levich from the Annual Meeting of the International Society of Electro-chemical Scientists in Zurich, are cases in point.

'Active' dissidents and refusniks con-

stitute, however, only a small part of the Soviet scientific community—indeed, a main campaigning point for refusniks is that, since they are by no means irreplaceable, the Soviet Union can well afford to let them go. (Thus, such vocal and embarassing dissidents, as Chalidze, Zhores Medvedev, and mathematician Aleksandr Esenin-Volpin, have been forced into exile.) So what of the bulk of the scientific community?

The overwhelming majority of scientists, it may be said, seem to find relief from discontent and frustration in 'inside emigration'—a complete absorption with work (although even that is a kind of dissent, since, according to *Pravda*, "a scientist's first duty is to be a patriot", with all that patriotism demands in voluntary community service and attendance at meetings and hortatory or celebratory rallies). The silent majority is, as everywhere, completely silent. Nevertheless, the scientific establishment still preserves some vestiges of independence from the State. The Academy of Sciences is the only remaining body in the Soviet Union to exercise the right of elections by secret ballot without official intervention (though it must be admitted that its current President, Aleksandrov, was nominated by the Party, before being duly elected). In spite of periodic press campaigns against Academician Andrei Sakharov, and in spite of the refusnik status of Corresponding-Academician Levich, the Academy has not, so far, expelled either. A two-thirds majority is needed for expulsion, and the lack of action so far indicates a degree of uncertainty as to how far dissidence, or fellow-feeling for dissidents, might spread.

In this situation of frustration, it is small wonder that so many appeals, whether from dissidents or refusniks, call upon the world scientific community to take some action on behalf of harassed Soviet colleagues. The appeals are wide-ranging. Scientists in the non-communist world may be urged to plead for clemency for victimised Soviet scientists on trial. They may be urged to press for reviews of particularly harsh verdicts against those already convicted. Or they may simply be asked to request visas for specialists wishing to attend international conferences. A recent statement from the US Committee of Concerned Scientists stressed that these actions are more than gestures of common humanity towards distressed colleagues. When a scientist is for any reason prevented from working, the research he might have done is "irrevocably lost to science"; the fate of dissident and refusnik scientists affects not only Soviet science, but also the whole worldwide scientific community. □

Soviet dissidents (2)

Keeping the flame alight

Robert Adelstein participated in the Moscow Seminar earlier this year. Here he describes his experience

"THE West is serving as a witness to our scientific death. This is the whole point of the Soviet policy, to doom us forever as scientists". The speaker was Mark Azbel, head of the best known Moscow Seminar of dissident scientists. The words were the emotional highpoint of an impassioned closing speech delivered earlier this year during an extraordinary session of the seminar. For five days Hershel Markovitz, a Professor of mechanics and polymer science, and I, a biochemist, shared in the science, thoughts and experiences of the seminar members.

We had arrived in Moscow six days earlier not only to participate in the Moscow Seminar, but more important, to express the continuing interest of western scientists in their less fortunate colleagues. Any prior doubts we had about the impact of a visit by two relatively unknown scientists were quickly dispelled by the overwhelming reception we received. In honour of our arrival it was decided to hold a 'symposium' during which we would deliver three papers each and the Russian scientists would contribute ten of their own. We would meet every day for 3–4 hours of formal science followed by informal discussions about the status and problems of dissidents.

The Moscow Seminar on Collective Phenomenon, as it is officially known, convenes every Sunday at 12 noon in the sparsely furnished apartment of Mark Azbel. (It is one of six different seminars of dissident scientists convening in Moscow at present.) It comprises approximately 30 scientists drawn from various disciplines: physics, mathematics, cybernetics, electrochemistry, biophysics and molecular biology. But despite their differences in training the scientists all share one common attribute—they have applied for an exit visa to Israel and the visa has been refused. They are refusniks.

With very few exceptions refusniks are fired from their jobs and forced to eke out a living either by tutoring privately or by securing menial, non-scientific work. They are denied access not only to laboratories, but also to libraries, where their names are removed from published works, and where the books they have written are removed from the shelves. Reference by Soviet scientists to their published works is forbidden. They are denied official permission to publish scientific

papers in the Soviet Union or to mail them abroad to foreign journals (though papers can, of course, be smuggled out). Incoming mail, particularly from abroad, is interrupted. Telephones are often disconnected. Attendance at scientific meetings at home or abroad is forbidden. And refusniks are often ostracised by former colleagues and harrassed by the KGB. The price of applying for an exit visa is high enough to discourage many would be emigrants.

Yet occasionally, in a pattern that defies discernment, a refusnik is released. My first visit to the Moscow Seminar had been in August 1972, the year it was first organised by 15 scientists because of "an urgent fear that we would lose our scientific standing, together with the loss of our employment". The words were spoken by Alexander Voronel, at whose apartment the seminar was convened. Three years after applying for his visa, Voronel was permitted to emigrate and is now a Professor of Physics at Tel Aviv University in Israel.

I remember sitting in Voronel's apartment with Benjamin Levich—a refusnik, a corresponding member of the Soviet Academy of Science, and a world-renowned electrochemist—together with his two sons Alexi and Yevgeny. His sons now live in Israel—though Yevgeny, a physicist, was incarcerated for a year in an Arctic Circle prison camp before being granted a visa. But their father and mother still await visas, despite Soviet promises to release them in October 1975.

My latest visit, in late winter, started on a Friday in a Moscow still blanketed with snow. My arrival with Hershel Markowitz, earlier in the day, was uneventful except for an extremely thorough search of our luggage at customs. Interestingly, the only reading matter which aroused enough curiosity to require further perusal by a higher official was not a month's supply of *Nature* and *Science* along with assorted scientific texts, but a Moscow guide book published in 1974 in the USA. It was our sole encounter with Soviet officialdom.

Later that night I journeyed by subway to the apartment of Irene and Victor Brailowsky—she a mathematician and he a cyberneticist—both of whom are members of the seminar. Arrangements were made to notify the seminar of our arrival (no easy task

Victor Brailowsky, harassed and threatened

since telephone communication was not possible) and to meet on Sunday for the trip to Azbel's apartment, which is located on another side of the sprawling city.

Mark Azbel greeted us enthusiastically on our arrival at his apartment. The balding, red-haired leader of the group looks considerably older than his 43 years. His whole manner reflects warmth and a certain dynamic intensity which seems to inspire the group whether discussing science or politics. After our introduction to the group—most in their 30s and 40s and much younger than I had imagined, we were joined by Benjamin Levich. A schedule for the 'symposium' was quickly drawn up, it being agreed to meet in Azbel's apartment every evening at six o'clock with the exception of Wednesday, when we would meet at Levich's. This would allow us to meet with some physical chemists who only attended Levich's Wednesday seminar, and with Alexander Lerner, like Brailowsky a refusnik cyberneticist. In one of the best examples of how international cooperation among scientists can force the Soviets to alter their policy, Lerner was permitted to attend an international meeting in Tbilisi, USSR, in August 1975, after western scientists threatened to boycott the conference.

The seminar on Sunday was brought to order by Azbel, raising his voice in a manner that brought to mind Moses at the Red Sea: "Yehudim Sheket"—literally, "Jews be quiet" (I have the impression that those are the only two Hebrew words he knows). Together, we covered a wide range of topics. Hershel Markowitz dealt with polymer rheology as well as with linear and non-linear viscoelasticity. I dealt with contractile proteins in muscle and non-muscle cells. Azbel discussed the decoding of DNA; Levich, charge transfer reactions in

solution; Victor Brailowsky, an algorithmic approach to diagnostic and prognostic medicine; Felix Lerner, electromicroscopy of muscle phosphorylase; and Edward Trifinov some biological consequences of clustering of pyrimidine photodemers in DNA. All of the lectures were presented in English, and translated into Russian for five or six seminar members who did not understand.

Trifinov, a 39-year-old survivor of the World War II blockade of Leningrad, whose father died in a Stalin camp, teaches his own course in molecular biology attended by the children of dissident scientists, among others. His first lecture coincided with the third day of the symposium, and he was terribly excited when I presented him with my last minute purchase, a new edition of Watson's *Molecular Biology of the Gene.*

At the conclusion of our first lecture, Azbel presented Hershel Markowitz and me with commemorative pins labelled with the initials MSCP—Moscow Seminar on Collective Phenomenon. The pins, awarded to all seminar lecturers, contain two lines symbolising uncoiled DNA and a balance—which Azbel pointed out as the "scales of justice".

Each evening when the scientific seminar was concluded, we gathered around a table for a snack, usually consisting of bread, cheese and fish, and the informal seminar commenced. Next to talking science, we spent most of our time discussing the plight of the dissidents, particularly why they had applied to leave, and what scientists in other countries could do to help. The two most frequently cited reasons for leaving were antisemitism and the corruption of scientific research by politics.

According to the members of the seminar, anti-semitism has become a sanctioned policy of the Soviet Government, manifested in the systematic exclusion of Jews from the better universities such as Moscow State. Moreover, Jews are being denied access to many professions so effectively that Alexander Lerner, who before his dismissal occupied an extremely important position, told me he applied to leave because "there is no future here for my children".

As scientists, Jews find the situation in the Soviet Union intolerable because scientific research is so enmeshed in politics. Promotion, or even survival in one position, requires constant political clearance, which in turn requires scientists to advocate political positions and sign political statements that they completely disagree with or know to be patently false.

So how can other scientists help? The seminar members stressed several possibilities. First, continued pressure must be maintained on the Soviet auth-

orities to allow free emigration; second western scientists should refuse to deal with Soviet scientists known to be instrumental in having dissident colleagues fired (a partial list of names is available); third, scientists, particularly members of official delegations, should visit the seminar and other dissident scientists (such visits, in the opinion of the seminar members, are instrumental in protecting dissident scientists from harassment by the KGB).

After meeting and eating with members of the Moscow Seminar for five days, Hershel Markowitz and I came away with the impression that they are indeed dying from a professional point of view. This isolation, as pointed out by Benjamin Levich, is particularly devastating for the younger scientists, whose careers and development demand constant work and access to recent publications. (Some of the work presented to us at the seminar was carried in 1971.) But despite their dire situation, they remain committed to their cause, and hopeful that help will come from scientific colleagues in other countries.

At an impromptu banquet held on the night before our departure, Hershel Markowitz and I augmented the usual diet with a chocolate cake, the lucky result of joining one of the many lines that suddenly form on Moscow sidewalks. The 'party' was tempered for me by two things. The first was a feeling of unreality brought on by the clash between a strong sense of identification with my dissident friends on Thursday night and the knowledge that on Friday afternoon I would be able to stop by my own laboratory to check on the latest results.

The second was Azbel's closing address. Besides warning of the potential scientific demise of dissidents, he pointed out that "although most Russian scientists are resigned to restrictions on their scientific freedom, for the dissident scientists this situation is intolerable. The dissidents would rather trade their prestige and, if necessary, their careers, for freedom. In this sense all scientists who apply to leave the Soviet Union must be understood as fighters for the freedom of conscience".

● Six months after my departure from Moscow little has happened to alter the status of the seminar members. Benjamin Fain, one of the few seminar scientists still working during my visit, has been fired. Several other members, including Azbel, Levich and Viktor Brailowsky have been harassed and threatened by the KGB, in an effort to get the seminar to disband. But still the meetings continue, soliciting visits from western scientists, among them the members of official delegations.

One month ago I received a letter smuggled out of the Soviet Union from Viktor Brailowsky. He and his two children have been granted permission to leave, but as often occurs there is a catch. Irene, his wife, is still refused permission—and the family will not leave without her. In his letter he appealed for scientists to send "letters, telegrams and telephone calls (the last are helpful) to R. V. Khochlov, Rector of Moscow State University", asking that she be permitted to emigrate together with her family.

Enclosed with Brailowsky's letter was a letter from another refusnik, German A. Shapiro, an endocrinologist and a former Assistant Professor at the Institute of Biophysics in Moscow. Discharged after applying for a visa in 1972, he now works as an ambulance attendant. The official reason for denying Shapiro permission to emigrate is that he had treated "patients who might know state secrets".

Two incidents illustrate that what at times appears to be a ludicrous situation, actually involves life and death. Tanya Levich has recently suffered a heart attack. Three years ago when the Soviet Government sent her son, Yevgeny, to a prison camp in the Arctic Circle, one of the reasons given for not letting Benjamin Levich emigrate was that it would be cruel to break up the family (since Yevgeny was obviously not free to leave the country). Now, having allowed Yevgeny and his brother out, the Soviets seem intent on making sure that Benjamin and Tanya never see their sons (and grandchildren) again.

The second incident involves another letter delivered by 'special courier'. While I was in Moscow, Alexander Lerner, knowing I had training in cardiology, showed me some electrocardiograms taken on Yefim Davidovitch. Davidovitch was a much decorated World War II hero who chose to sacrifice his career in an effort to emigrate to Israel. He was, however, quite ill, and clearly had heart disease. His friends, including Lerner, were concerned that he might not survive the difficult life imposed on refusniks (his phone had been disconnected and he had been refused treatment by certain doctors and hospitals). After inspecting the cardiogram, I suggested that letters to some high Soviet health officials might be helpful in gaining his release on medical and humane grounds. I asked that a copy of his medical record be forwarded to my laboratory so that I and others could write on my return. Approximately two months after I returned, a detailed report on his medical history and physical examination arrived. Unfortunately, Yefim Davidovitch had died two weeks earlier. □

Soviet dissidents (3)

View from the Promised Land

Nechemia Meyers assesses the opportunities for
Soviet scientists arriving in Israel

Alexandr Voronel, vantage point

"AMONG the Russian immigrants are scientists on a par with the best in the US, and their coming to Israel gives this country a unique opportunity to develop many new research areas". So says Professor Alexandr Voronel, a Soviet expert in solid state physics formerly associated with the Institute of Physics at Chernogolovka, near Moscow, who joined the staff of Tel Aviv University. Voronel and his colleagues fear that this opportunity may be missed for lack of a coherent science policy in general and a plan for the absorption of these scientists in particular.

Exactly how many immigrant scientists have come here from the Soviet Union is impossible to determine, particularly because the term "scientist" covers a broader category of people in the East than it does in the West. Absorption authorities speak of 800 scientists and 400 graduate students (among the 100,000 Soviet immigrants who have arrived since 1971). Voronel thinks there are between 1,000 and 1,500, making up perhaps 15% of Israel's scientific community (as compared to about 25% of her doctors).

Those who have found employment in their professions work for the most part in universities and research centres. However, many hold temporary positions financed not by regular budgets but by special two- and three-year allocations from the Absorption Ministry and the Jewish Agency. When these allocations run out, the institutions where they are now working, which are so short of funds that they have had to fire some members of their academic staffs, will find it difficult to offer the immigrants permanent posts. In any case, future immigrants will find the universities closed to them.

This being the case, efforts are being made to channel immigrant scientists into industry. For example, the Centre for Absorption in Science announced that it was now sending 40% of its applicants to industrial firms, as compared to 10% 2½ years ago. However, as things stand, the absorptive capacity of Israel industry is also limited unless structural changes take place. For their part, the Russian immigrant scientists, or at least some of them, feel that Israel should pay heed not so much to the American experience, as some have suggested, but to the Russian one.

"When I was in the Soviet Union," Professor Voronel stated earlier this year, "I was surprised that the Americans with all their resources and techniques had no scientific advantage over the Russians in many fields. Now that I am in Israel—that is to say within the framework of Western science—I wonder why the Russians are not doing even better, because from my vantage point I can see great advantages to the organized and planned system that exists in the USSR." Voronel says that with rare exceptions, such as the Manhattan Project or the attempt to land a man on the Moon, US research is based on individual initiative and achievements. In contrast, the USSR creates scientific teams geared to achieving specific national goals.

Voronel and his colleagues fully appreciate the freedom offered to them in Israel, which they would not forego in any circumstances; but they nevertheless believe that Israel requires a less "American approach" to science policy. "The US is big enough," Voronel declares, "to achieve success, without overall planning, in a great many fields. Israel, in contrast, will achieve very little without concentrating its efforts on a few well-defined projects, and this is where immigrant scientists from the Soviet Union can be of great value. They know how to organise and apply research though, of course, they must be better acquainted with the requirements of the Israeli economy before they are in a position to make concrete suggestions."

The Diamant brothers, Lev and Emanuel, are also drawing on the experience of Russian science in their attempt to create Arshach, a science-based community in the Galilee. They hope it will eventually expand into a kind of Academ Gorodok, the large Siberian research centre near Novosibirsk where Lev Diamant studied and worked in plasma physics before coming here. In spite of the fact that both brothers found positions in Israel soon after their arrival—Emanuel Diamant as an engineer at Tel Aviv University and Dr Lev Diamant as a physicist at the Weizmann Institute—they wanted to do more than just earn a living. They wanted to lay the foundations for the absorption of hundreds or even thousands of fellow scientists from the Soviet Union in a self-supporting scientific community to include both research and development and production facilities.

Their little settlement, perched on the Galilean hills not far from the spot where Christ gave his Sermon on the Mount, is no Academ Gorodok as yet. In fact, a lack of housing limits it to a few families and a modest research and development programme, most of it based on subcontracts from established research centres in such fields as solar energy, custom-made electronics, electro-optics and lasers. Lev Diamant, who turned down US job offers, is not worried that the settlement is growing very slowly or that it is some distance from existing scientific and industrial complexes where computers, libraries and other important facilities are to be found. "In the US many scientists travel 60 miles, for example between Burlington and MIT, without thinking about it; why then," he asks, "should we Israelis make such a fuss about travelling 90 kilometers between Arshach and Haifa?"

The settlement has already established itself as an intellectual centre for immigrant scientists from the Soviet Union. The famous scientific seminars held in Moscow by Jewish scientists who had been dismissed from their posts after applying for an exit visa have been recreated at the village. They are headed by Professor Voronel, who was himself a key participant in the Moscow seminars before finally being permitted to leave for Israel last year.

Yet the purpose of these seminars, is different. "In Moscow," Voronel explains, "we were forced to concentrate on theoretical subjects because the authorities had taken away our laboratories and our instruments. Here we are discussing applied research, which we are trying, so far without too much success, to develop in Israel."

Soviet dissidents (4)

Trying to keep in touch

Yevgeny Levich describes the complex procedure involved in obtaining permits for travel abroad

WESTERN scientists welcomed the development of scientific contact with the USSR which came with détente. But enormous obstacles remain. A Soviet scientist intending to go overseas to take part in a conference or in some scientific exchange programme has to go through a complicated bureaucratic procedure, like any other Soviet citizen going abroad.

The first stage consists in obtaining a character reference (*kharakteristika*) which must be confirmed by the administration of his place of employment, by the local party organisation and his trade-union branch. The reference deals, among other things, with such matters as the applicant's participation in public activities, his moral image and his relations with the all grades of staff at his place of employment. Accompanying the reference is a detailed questionnaire covering the applicant's nationality, his party membership, details about all his former places of work and so on. The applicant is also asked to supply detailed information about his close relatives (including divorced spouse). When he has obtained the reference he is called for an interview to the Regional Committee of the Communist Party, where he is given the instructions which complete the open part of the precedure.

The main mechanism of issuing permits for foreign travel remains hidden from the eyes of the applicants. The outsized apparatus of the foreign Section of the Academy of Sciences of the USSR (General Korneev of the KGB has been head of this section for many years) or the ministries in charge of the various scientific research institutes are the bodies which deal with "travel matters". The central KGB has the last word in making decisions, which are communicated to the prospective traveller, as a rule, a day or two before his intended date of departure. There have been cases when a travel permit was taken away when the person was boarding the aeroplane.

In practice, the administration of the local institute (and especially its "first section"—the secret section which exists in every scientific institution and is connected with the KGB) plays an important part in making the decision. The deputy director of the institution, who is normally the head of the "first section", is also responsible for all international contacts. The information about the applicant which is passed on

through the closed channels has much more weight than his official reference. After the sanction of the Foreign Section has been received and confirmed by the KGB the "travel case" is transferred to the "Commission on Travel", headed for many years by the Politbureau's chief ideologue Mikhail Suslov.

The considerable efforts which are made to obtain permission to go abroad reflect the attraction that travel has for the Soviet scientist. Purely scientific considerations play the significant part here, of course, especially considering the almost complete isolation of the Soviet scientists in the recent past; but such academic interests are reinforced by financial attractions and considerations of prestige. Foreign travel is graded, like the Soviet elite's other privileges, according to value. Short trips to scientific conferences for which the scientist himself must pay are thought to be least attractive. It is much more difficult to be included in an official delegation going abroad at the state's expense. Then there is the rare possibility of going abroad for a long period of time, for example, within the framework of scientific-exchange programmes. Additional grading differentiates between travelling to the countries of Eastern Europe (not always easily obtainable) and travelling to the West.

The possibility of receiving a permit to travel abroad, and the efforts which must be undertaken to obtain it, are determined by a vague but identifiable correlation between the prestige value of the trip and the applicant's personal qualities. Being a party member, holding an official academic position, having personal connections in the administration and personal initiative are all to the applicant's credit; having relatives abroad, an unsuitable nationality, and a lack of interest in the party are not. Negative characteristics are usually critical and they automatically block off the entrance to some—if not all—ladders of privilege. On the other hand, some persons can arise to a level of the administration and party where they can by-pass the usual procedures and plan their trips more freely.

Moving up on the ladder of privileges and, in particular, gaining access to overseas contacts presents additional demands, the violation of which could mean the immediate loss of all that had been achieved. Scientific contacts between Soviet scientists and the West

Sakharov, attacked

are subject to fairly strict control. Two examples:

● Every scientist must present a detailed report, including not only details of scientific and industrial significance, but also characteristics of all the Western scientists whom he met, to the Overseas Section of the Academy of Sciences.

● Each scientist meeting with overseas colleagues at his home or at his institute must receive permission for such a meeting from the "first section" of his work-place. He must also present a detailed report about the meeting to the "first section" afterwards.

Evasion of these obligations automatically stops access to foreign contacts. On the other hand, some Soviet scientists deliberately overestimate the importance of their overseas contacts and their knowledge of scientific and industrial works in the West in order to increase their chances of a permit for their next trip.

Strict control applies not only to personal contacts, but also to scientific publications: all materials sent abroad, including letters to colleagues, must go through the special commission in every institute, which include representatives of the "first section". After being checked by the commission the materials must receive the sanction of the censorship office (the *Glavlit*). A particular form of control has appeared with the inclusion of Soviet representatives on the editorial boards of some international magazines, providing an opportunity to prevent publication of articles in foreign magazines.

Those who can travel abroad because of their position naturally enjoy significant advantages in the struggle for power which goes on in the upper layers of the Soviet academic community. But whereas scientific life inside the USSR is controlled to a significant extent by the representatives of the Soviet scientific elite, in matters connected with foreign contacts the

critical influence is concentrated in the hands of the party apparatus and the KGB. The selectively chosen representatives of the scientific elite have, however, merged with the administrative and party elite: those who have signed letters attacking Sakharov, those who have dismissed Jewish scientists wanting to emigrate from the USSR, and later signed false secrecy certificates of these scientists, are themselves scientists. Their actions are dictated not by their scientific but by their administrative interests, by their proximity with the party oligarchy and their desire to use this in their struggle for rank and power. They are the ones who appear most often at international conferences; they are the ones who gain most from the "scientific détente". The strengthening of scientific contacts

with the West, as long as they develop in a way suitable for the Soviet authorities, increases the influence of the most rigid elements of the Soviet establishment and provides them with additional levers to influence scientific life in the Soviet Union.

Soviet citizens differ from one another in everything, just as citizens of any Western country do. In accordance with the spirit of the "third basket" of the Helsinki Declaration, contacts ought to be developed on the individual level and not on the level of the delegations selected by the authorities. The organisation committees of international conferences should reserve the right to invite to conferences Soviet scientists who are prominent in their field, and should make their attendance a condition for the attendance of the

official Soviet delegation. The participation of Western scientists in any scientific exchange programmes must be conditional on the Western side being able to choose at least some of the members of the Soviet delegation and being able to have free contacts with any scientists in the USSR whom they would like to meet. Soviet scientists must also retain the right to send their articles to foreign magazines through members of the editorial board as well as through the Soviet regional editor.

Measures such as these would not cause a halt to scientific contact with the West. The needs of the state, and of the members of the elite, ensure that. And they represent something more than mere declarations of support or solidarity. ☐

COMECON

WITH growing industrialisation, fuel and power supplies are a matter of increasing importance to the Comecon block, particularly as the new policy of economic integration begins to take effect, and considerable reliance has come to be placed on oil, gas, and electricity delivered from the Soviet Union by pipeline and cable. But despite official Soviet statements that fuel and power export commitments would remain unaffected by the new drive for economies at home, it seems that there will after all be an export cutback. During a recent Miners' Day rally in northern Bohemia it became apparent that deliveries of Soviet oil to Czechoslovakia over the next five years would "fall short" of original estimates. Accordingly, the Czechoslovak Government has concluded that "we shall be unable to maintain the growth rate of past years".

There are other problems too. An article in the Slovak *Pravda* notes that in spite of the additional 4,000 MW or so of new generating capacity planned for the next five years (which should include two atomic reactors at Jaslovske Bohunice, providing some 35% of the total new capacity for Czechoslovakia), the generating system will still have "minimal reserves" and will be constantly stretched.

Nuclear power is a matter of pressing importance to Comecon, and, not surprisingly, there is considerable international cooperation. Joint research into the construction of the necessary equipment, which has resulted in the development of new types of generator, is carried out by Interatomenergo and Interatomininstrument. The latter organisation

is the only Comecon enterprise to have its own autonomous accounting system and operating capital of convertible currencies. Atomic power stations are being constructed (with Soviet aid and expertise) throughout the Comecon block, notably at Lake Zarnowiec in Poland, and Paks in Hungary. Ultimately, they should all

be connected with a 750 kV grid, the first line of which, from Ukraine to Hungary (with a later extension planned to Yugoslavia), should be commissioned in 1978. By 1990, nuclear power stations should meet one-quarter of the forecast electricity demand of the Comecon block.

In the meantime, to fill the energy gap, conventional fossil fuels are in ever increasing demand, and intensive work is being carried out both to modernise existing coal mines and to develop new ones. One such project, described recently by the Prague television service, bears a touch of inadvertent humour. A new open-cast coal mine, which will provide fuel for a power station in northern Bohemia, is, in line with the Socialist custom of naming important enterprises after

revolutionary heroes, to be called after Maksim Gor'kii. One of Gor'kii's most famous works is entitled *'From the Lower Depths'*.

● The drought which this year has affected all the European members of Comecon together with certain western parts of the Soviet Union (though not the Moscow region, where too much rather than too little rain has been the problem), has forced the authorities to place a new emphasis on the utilisation of water resources. Some research had, indeed, already begun with the mapping last winter and spring of Carpathian water-resources. That survey did not, however, extend to the eastern regions of the Carpathians (Ukrainian SSR), where over-zealous tree-felling is apparently resulting in erosion and consequent "destruction" of the water table. Precipitation in the Carpathians is of vital importance to the water supplies of Poland, Czechoslovakia, Romania, Hungary and Ukraine.

Several local water economy schemes are now being implemented. These include anti-flood measures on the Tisza, Mura and Raba in Hungary, a dam on the Cirocha in Slovakia, a special turbine-type aerator for the purification of polluted river water (developed at the Purification Equipment Institute in Sofia) and a reservoir on the Odra near Mietkow in Poland. Although, like the Carpathian survey, many of these projects were already envisaged or under way before the drought became a pressing problem, the considerable media coverage they have received indicates that water supply has become an issue of grave importance.

IN BRIEF

Synthetic fuel block

A controversial plan to pump billions of dollars of federal funds into the commercial production of synthetic fuels in the United States was unexpectedly killed off in the House of Representatives last week. By the narrow margin of 193 votes to 192, the House blocked further consideration of a Bill to provide up to $4,000 million for the construction of large-scale plants to produce oil from shale, and oil and gas from coal. The vote means that there is now no chance that Congress will pass the legislation before it breaks up for the November elections. The Bill was a major plank in the Ford Administration's long-term energy strategy, which calls for the production of large quantities of synthetic fuels by the mid-1990s to replace dwindling domestic supplies of oil and gas. It was defeated by an odd assortment of bedfellows ranging from environmentalists fearing the pollution threat from conversion plants, to right-wingers opposing federal subsidies for industry.

Energy get-together?

Democratic Presidential candidate Jimmy Carter last week offered a plan to consolidate an assortment of government agencies in the United States into a single Department of Energy. Stating that "our country still has no energy policy" three years after the 1973 oil embargo, Carter said that the new department would include the Energy Research and Development Administration, the Federal Energy Administration, parts of the departments of Commerce and Interior, together with the Federal Power Commission. The plan would have a tough time getting through Congress if Carter is ever in a position to propose it formally, however. Former President Nixon twice suggested the setting up of a Cabinet-level department of energy and natural resources, but the reorganisation was never approved by Congress because it would have cut across the jurisdictions of many Congressional committees, which now roughly match those of existing departments and agencies.

Bangladesh council

The Government of Bangladesh has constituted a National Council for Science and Technology to plan the country's scientific and and technological programme and activities. The high powered council is chaired by the President of Bangladesh, and one of the two vice-chairman is an eminent scientist who is also the Science Adviser to the President, Professor M. Innas Ali.

The council has 19 other members who are *ex-officio* chairmen of research councils, secretaries of some ministries, Vice-Chancellors of two technological universities and two reputable Bangladesh scientists or technologists nominated by the President. Professor Ali himself has a distinguished career in scientific research, education and administration. He holds a Master's degree in electric engineering from the University of New York, a doctorate in nuclear physics from the University of London; and is a former Vice-Chancellor of Chittagong University.

SECRECY, in science as in politics, is a delicate and difficult subject. There is much concern when the details of Cabinet deliberation are prematurely revealed, yet at the same time there is a strong campaign for the relaxation and rationalisation of the Official Secrets Act. It is difficult to strike a happy medium. The results of scientific research concerned with military problems or commercial activity may legitimately be kept secret for a time, but it often seems that publication of much of this work is delayed unnecessarily. Some results which might have a wide application never see the light of day, even when the defence of the country is no longer at risk, and when no benefits would immediately accrue to a company's commercial rivals.

The reputation of a scientist depends largely on the quality and quantity of his published work. If he is unable to publish, he is likely to remain unknown except among his immediate colleagues. It is therefore important to the individual, as well as to the scientific community, that as few restrictions as possible be placed upon the publication of the results of any kind of research. High salaries, or even the award of honours and decorations to otherwise unknown scientists may be poor recompense for the denial of proper recognition within the scientific community.

When the reorganisation of

government-supported research attributed to Lord Rothschild was being discussed, there was some concern that it might have a harmful effect on the dissemination of scientific know-

---

In confidence

KENNETH MELLANBY

---

ledge. It was feared that the executive government departments, to whom so much of the research funds was to be transferred—so that they, as "customers", could place "contracts" with research councils or other laboratories—might be reluctant to allow the scientists who did the work to publish their results in the accepted scientific journals. These particular

fears appear to have been unfounded. Some senior ministry officials have suggested that since they paid for the research, they should have the right to decide what should and should not be published. But it is to the credit of Chief Scientists that they normally insist that those engaged in research should be free to publish as they wish. Indeed, in some cases contracts have included the proviso that results should be made available to the customer first, and there seem to have been few cases where final publication has been seriously delayed.

And yet the new system appears to have reduced substantially the amount of material being published by some of the laboratories now largely engaged in contract research. The reason would seem to lie in the nature of the contracts themselves. Few contractors seem to have followed the example of the Ministry of Agriculture, Fisheries and Food, whose contracts have usually been for very large sums, and have covered substantial fields of study, so as to allow room for manoeuvre and originality. Too many contracts have been small and circumscribed by limited objectives, presenting few opportunities for research workers to make significant advances. There has consequently been little worth publishing. And in that situation, it does not matter whether or not the workers are pledged to secrecy.

news and views

Complementary mispairs

from E. G. Richards

THE fidelity of DNA replication is a matter of crucial importance : too low, and life as we know it is impossible; too high and there is no evolution for us to ponder the matter. The observed frequency of mutations and replication errors must depend on several factors, including the nature of the error, the chemistry of the replication machinery and the fundamental chemical properties of the bases in DNA themselves.

An interesting theory has been propounded by Topal and Fresco (*Nature*, **263**, 285; 1976) to account quantitatively for the observed frequencies of a certain class of mutation—point mutations which entail the substitution of one base pair for another in the DNA.

Such point mutations can be of two sorts: transversions in which an A–T or G–C base pair is substituted for its opposite number, namely a G–C or A–T pair respectively; and transitions in which the pair is reversed between the two DNA strands so that an A–T pair becomes T–A or G–C becomes C–G.

As is often stated, the Watson–Crick base pairs, A–T and G–C, are each complementary. Among other things, this implies that both sorts of base pair are geometrically similar and can substitute for each other in the regular DNA helix. The crucial observation, based on model bulding, of Topal and Fresco is that there are other sorts of base pair which are geometrically similar to the normal Watson–Crick pairs. These they have termed complementary mispairs, and they are formed from bases which are in thermodynamically rare forms such as the enol or imino forms or in the *syn* rather than the normal *anti*-configuration (obtained by rotating the base through 180° about its glycosidic bond).

In the replication process, each base in turn selects, with the aid of the enzymes involved, a complementary partner which is incorporated into the growing chain. If on rare occasions a base, instead of selecting its normal partner to form a normal Watson–Crick base pair, selected another partner to form a complementary mispair, an error or mutation would be born. Topal and Fresco find that all possible transversion and transition mutations could come about in this way.

Considerable credence is lent to this idea by the report that base oppositions corresponding to each of the required complementary mispairs have been demonstrated experimentally to take up intrahelical conformations in long double stranded polynucleotides dominated by normal Watson–Crick base pairs, whereas other base pairs which are not complementary and are not required take up extrahelical conformations instead.

So far one might suppose that the frequency with which a transversion or transition occurred would be determined mainly by the proportions with which the different nucleotides exist in their rarer tautomeric or isomeric forms. These proportions are determined by the thermodynamic properties of the bases and have in fact been measured. The results suggest that the enol or amino forms would occur at a frequency of 1 in 10^4–10^5 and the *syn* form at 1 in 10–20. These frequencies are however far higher than those at which point mutations are observed to occur—1 in 10^9–10^{12} base pairs replicated.

In circumventing this difficulty, Topal and Fresco note that there is good experimental evidence that replication takes place in two stages. In stage one, the replicative machinery holds the template base in position and presents a hole into which only a complementary base may fit to be joined to the growing strand. At this stage a mispair may be formed with the frequencies noted above. In the second stage a checking mechanism is brought into play which ascertains if the newly formed pair has standard geometry. If not it is excised. Topal and Fresco propose that before the check is completed, thermodynamic equilibrium is re-established and the tautomers revert to the keto or amino form with a corresponding disruption of the geometry (because of stacking interactions it is likely that the *syn-anti* equilibrium is frozen). At frequencies determined by the same thermodynamic considerations, however, a small proportion of mispairs will remain mispaired and not be detected. Thus the proportion of mispairs that are both formed and survive the checking step will be approximately the square of the proportions mentioned above. Topal and Fresco have worked out in detail from their proposed mispairs and the relevant thermodynamic data the frequencies at which they would expect the various sorts of point mutation to occur. They find a remarkable agreement with data obtained from a study of the frequency of spontaneous mutations in the tryptophan synthetase A gene of *Escherichia coli*.

One interesting feature in the proposed base mispairing schemes is that transitions arise only from purine–pyrimidine mispairs and transversions only from purine–purine mispairs. It is therefore of interest and wholly consistent with Topal and Fresco's hypothesis that the base analogue 5-BrU has the specific effect of enhancing the frequency of transitions. 5-BrU also has a much greater tendency to enolise than its natural counterpart T. Similarly they can explain the specific effect of

2-AP in inducing trasversions. They go on to comment that chemical mutagens (many of which are carcinogenic) may act by disturbing the tautomeric or isomeric equilibria of bases with which they interact.

This hypothesis is certainly elegant but as the authors point out, further experiments are required to assure its validity. Be this as it may, if it is valid, similar effects may well occur in other situations in which complementary base pairing is invoked. In particular it may occur in transcription and in mRNA interactions with tRNA during translation on the ribosome. Such matters are the subject of a second paper by Topal and Fresco (*Nature*, **263**, 289; 1976).

Any discussion of codon–anticodon interactions in translation must be dominated by Crick's wobble hypothesis. Topal and Fresco devote part of their paper to a discussion of what base pairs could and could not occur in the third codon wobble position. One suggestion they come up with is that in the wobble pair A–I, the A residue on the codon may be in the

syn configuration rather than the *anti* as proposed by Crick. This would give a glycosidic bond separation close to the standard 10.9 Å instead of the longer value of 12.8 Å.

Whatever the precise nature of the wobble pairs may be, mispairing would lead to errors in amino acid incorporation. The *in vivo* rate of such errors is as high as 1 in 10^4. This figure, Topal and Fresco surmise, results from mispairing in the transcription process as well as in all three positions in the codon–anticodon interaction. The calculation of the value expected on the basis of mispairing schemes and the tautomeric equilibrium constants is complicated by the degeneracy of the code, but ignoring this difficulty, Topal and Fresco arrive at a value of 1 in 4×10^4 for the overall error rate. Presumably there is no checking process analogous to that in replication. This may be because the errors can arise at two distinct stages—transcription and translation—at roughly the same frequency so that the economics of the cell would render it unprofitable to check errors at one stage only. ☐

climates to observe the shape of a single flagellum, even though its diameter is only 20 nm. The advantages of this over electron microscopy are that distortions and environmental changes induced by specimen preparation can be avoided, and phase transitions observed directly while the medium is changed. Such transitions include a change from left- to right-handed helices on reducing the pH.

Kamiya and Asakura describe how this change occurs in reconstituted *Salmonella* flagella. The change is rapid and reversible, and can be seen to start at one end of a flagellum and propagate down its length. As this happens, the transformed region appears to rotate rapidly round the remainder, which in their experiments was tethered to the slide. Similar effects can be induced by hydrodynamic forces, and Hotani's ciné film of this sort of motion has recently been gripping scientific audiences around the world. The experiments show clearly that the transformation is due only to a reorientation of the subunits, in a process akin to that postulated by C. R. Calladine (*Nature*, **255**, 121; 1975) to explain the different static waveforms that can occur.

Hotani's companion paper concentrates on the static appearance of flagella of mixed type, in which polymerisation is started with one monomer mixture and finished with another. As well as measuring accurately the parameters of the helices, Hotani measured the "block angle" between the axes of helices at the point of changeover. There seems to be no kink in the structure, but the transition from one form of packing to another causes the new helix to coil about a different axis. This is particularly marked when the handedness reverses; the block angle can then be as small as 94°. The most interesting feature of Hotani's results is that nearly all the data fit a simple equation relating the block angle to the pitch angles of the two helices concerned. This equation demands a condition which also underlies Calladine's prediction of helical polymorphs.

This prediction followed the observation that a flagellum is made up of longitudinal rows of subunits which generally make a small angle with the rod axis. The large scale helices adopted are consistent with some of these rows running always along the inside of the helix, while other rows (of slightly greater length per turn of helix) run always along the outside. Calladine showed how "short" rows could arise by rearrangement of the connections between subunits, and he argued on purely mechanical grounds that such rows should cluster on one side of the rod. A transition between polymorphs would then involve a

Structure of bacterial flagella

from Michael Spencer

BACTERIAL flagella fascinate biologists, molecular and otherwise, for a number of reasons. They are built up from identical subunits; they are helical (look what that did for Watson and Crick); and they constitute a largely unsolved problem in motility. It seemed a couple of years ago that the important question of whether they rotate bodily or merely execute helical waves had been settled in favour of the former idea; certain lines of indirect evidence convinced many scientists that there must be some kind of molecular motor, complete with bearings, at the base of each flagellum. However, direct evidence for the functioning of such a device has been lacking, and a minority remains sceptical. The case for the new doctrine was summarised in a persuasive article by Howard C. Berg entitled *How Bacteria Swim* (*Sci. Amer.*, 36, August, 1975).

Meanwhile a number of workers have refrained from committing themselves firmly to any hypothesis, and have continued to publish new structural data. The most prolific of these has been Sho Asakura of Nagoya University, who has specialised in the

repolymerisation of flagella *in vitro*. By mixing seeds and monomers from various mutants, a number of flagellar forms can be produced which vary in the pitch, amplitude and handedness of the large-scale helix into which the basic rod structure is coiled. Others have used optical diffraction of micrographs to study the detailed packing of subunits (for example, Kondoh and Yanagida, *J. molec. Biol.*, **96**, 641; 1975). It is unfortunate that almost every worker in the field has invented a new terminology to describe the helices, and evocative but imprecise terms such as "curly" and "semi-coiled" have served to keep the non-specialist confused.

The two most recent publications in this field are by Hotani now at Kyoto University, and by Kamiya and Asakura (*J. molec. Biol.*, **106**, 149 and 165; 1976). Both papers describe the use of dark-field light microscopy, which has recently undergone quite a renaissance. Thirty years ago it was used by Pijper to observe flagellar bundles in motion, using sunlight on the roof of his South African laboratory. By using a high-intensity lamp it is now possible for workers in shadier

change in the number of "short" rows, but not in the side on which they clustered. Hotani's results are consistent with this argument.

What is intriguing is that the kind of structure described is also required by the basic "non-rotary" hypothesis for generation of helical waves, an idea with a lineage going back to 1883. On this hypothesis the packing pattern would be induced by some mechanism at the base to propagate round the rod, giving rise to apparent rotation. Although Hotani follows Kamiya and Asakura in eschewing speculation, their results taken together suggest that there could yet be some life in the old idea. □

The folded chain's last stand?

from Paul Calvert

IN the past two years low angle neutron scattering studies on amorphous polymers have shown that the chain dimensions in both the liquid and glassy states are equal to the random coil dimensions seen in ideal solution. This scotched the idea that there are ordered regions in amorphous polymers but has left the puzzle of how to pack chains to high density yet retain the random coil (Kirste *et al., Polymer*, **16**, 120, 1975; Cotton *et al., Macromolecules*, **7**, 763, 1974; Wignall *et al., Eur. Polymer J.*, **10**, 861, 1974).

Results of similar experiments on crystalline polymers are now being published and very surprisingly show that the chain is still in a random coil. This one fact contradicts much of our current understanding of polymer crystallisation.

The small angle neutron scattering experiment is similar to small angle X-ray scattering except that the contrast arises from the very different scattering powers of perdeuterated and protonated polymer rather than from density differences. Usually a few percent of normal polymer is mixed into a matrix of wholly deuterated chains. The strong incoherent scattering from the protonated chains is measured as a function of angle and allows their radius of gyration and molecular weight to be determined. For amorphous polymers the radius of gyration was found to be equal to the ideal solution value. The apparatus necessary is a nuclear reactor with a thermal neutron beam and a diffractometer. Experiments have been done at Harwell but the beam intensity there is low and much more can be done on the purpose-built research reactors at Grenoble and Jülich

Was the early Solar System windswept?

from David W. Hughes

IF the nebula that formed the Solar System had the same elemental composition as the Sun, the condensates, that is the planets, asteroids and comets which make up our present Solar System, represent only about 5% (by mass) of the original nebula. The remaining 95% has been lost in the intervening time between planetary formation and the present. Most modern cosmogonical theories use the solar wind to sweep the newborn Solar System free of this "left over" nebula which is made up mainly of volatile elements such as hydrogen and helium. This solar wind is a cloud of particles (mainly protons) which are streaming away from the Sun. At the present time the average cloud density near the Earth is 5 protons cm^{-3} and the particles stream past at a speed of about 450 km s^{-1}. In the T Tauri stage of the Sun's evolution this wind was blowing at about the same speed but was much denser, the Sun losing mass at a rate which could have been as high as 10^{19} g s^{-1}. This T Tauri stage lasted around 50 million years during which time the Sun's radius decreased from nearly that of the radius of Mercury's orbit to its present 7×10^5 km and the Sun lost a considerable percentage of its mass. The Sun was also a variable, surrounded by a thick, highly active chromosphere. Gravitational collapse was giving way to thermonuclear hydrogen burning as the main energy source

M. J. Handbury and I. P. Williams (Queen Mary College, University of London) question the wind-sweeping

hypothesis in a recent paper in *The Observatory* (**96**, 140; 1976). Assuming that the solar gaseous nebula is in equilibrium about the Sun before the onset of the wind the authors then calculate the rate of momentum transfer between a continuous, spherically-symmetric wind and the nebula. Three forces are acting on the nebula, gravity, "centrifugal force" due to its rotation and the solar wind force due to its rotation and the solar wind force due to momentum transfer. As angular momentum is conserved a formula can be easily obtained for the new equilibrium distance between the nebula and the Sun. Handbury and Williams find that for all reasonable values of the parameters the nebula does not get pushed out to infinity by the T Tauri wind but simply increases in radius.

Now if energy transfer and not momentum transfer is the actual mechanism the total nebular energy could become positive and the nebula would disperse. But this would require the kinetic energy of the wind to be converted into thermal energy in the nebula and then back again into kinetic energy as the nebula moves away. Also loss due to radiation, evaporation or fast moving particles and other mechanisms has to be negligible.

So the authors conclude that the solar wind cannot be the broom which swept from the Solar System those elements that did not condense to form the planets and that cosmogonists must look for another process.

(see Schmatz *et al., J. appl. Cryst.*, **7**, 96, 1974; for a general description of small angle neutron scattering).

The first experiments on crystalline material showed that the scattering was dominated by scattering from the matrix attributed to submicron voids. This effect could be eliminated by using deuterated polymer in protonated matrix rather than *vice versa* as the scattering lengths of the protons and carbons almost cancel. It then became clear that segregation was taking place so that the deuterated polymer was clustering in the crystals as the measured molecular weights were many times the true value (Schelten *et al., Polymer*, **15**, 682, 1974: *Colloid and Polymer Sci.*, **252**, 749, 1974). This arises because deuterated polyethylene melts at about 6 °C below the protonated polymer.

The next set of results to be pub-

lished were those of Sadler and Keller (*Polymer*, **17**, 37; 1976). They measured solution grown single crystals in the range of scattering angle appropriate to the interchain spacing in the crystal rather than the chain as a whole. They showed that their results can be explained in terms of an adjacent re-entry model but could not really eliminate alternative models. These experiments may have been affected by clustering which could invalidate the results.

During this summer results on melt crystallised polymers have been reported at several conferences by Ballard, Wignall and Longman from ICI at Runcorn using the Jülich diffractometer (to be published in *Polymer*). They tried to eliminate clustering by using a branched chain protonated polyethylene in conjunction with a linear deuterated polymer so that the crystallisation temperatures

matched. They found that even with such matching they could not eliminate segregation except in samples crystallised by quenching. Neutron scattering measurements were made on these samples down to low enough angles to observe the radius of gyration of the whole chains. For quenched samples they found a value of ratio of radius of gyration to the square root of molecular weight $\langle S^2 \rangle_w^{\frac{1}{2}}/M_w^{\frac{1}{2}} = 0.46 \pm 0.05$, the same as for molten polymer and ideal (θ) solution. Over the molecular weight range from 40×10^3–400×10^3 it is not possible to tell whether there is a change in the ratio. At concentrations of deuterated polymer less than 10% the quenched samples gave correct molecular weight values showing that no clustering occurs. Slowly crystallised samples give a molecular weight value several times too high due to clustering.

The significance of these results is that they undermine the folded chain theory of crystallisation in which each polymer chain crystallises in a 'jumping jack' configuration, repeatedly traversing the crystal thickness of about 100 atoms (a stem) then folding back on itself. The chain will thus have the shape of a thin sheet. This model is supported by a large amount of indirect evidence on the morphology and properties of single crystals grown from solution and by the Lauritzen–Hoffman theory of crystal growth which correctly predicts the crystal growth rates from melt and solution on the basis of adjacent re-entry folding.

One direct experiment is that of Bank and Krimm (*J. Polymer Sci.*, **A-2, 7,** 1785; 1969) who studied the infrared spectra of mixed perdeuterated and normal polyethylene and showed that even at low concentrations a deuterated stem is likely to be adjacent to other deuterated stems so that the chain must fold back on itself. The Sadler and Keller neutron experiments are equivalent to this. Mandelkern and coworkers (*Macromolecules*, **4,** 672; 1971; reply, *ibid,* **5,** 209; 1972) have raised the question of segregation in the Bank and Krimm experiments but could not really prove it. The new neutron experiments seem to have verified Mandelkern's idea and invalidated the infrared experiments. The main alternative model to adjacent re-entry folding is the 'telephone switchboard' or non-adjacent re-entry proposed by Flory but a few months ago the majority of polymer scientists accepted adjacent re-entry as the norm.

Ballard and coworkers discuss their results in terms of two models. The first, assuming the chain is in a random coil configuration, gives a good fit to both the radius of gyration and the shape of the scattering function. If adjacent re-entry folding is assumed, a fit can be obtained if the chain folds a

specific number of times then moves to a new crystal, but the constraints on the number and length of folds are such that it becomes difficult to explain the molecular weight dependence of the data. Thus the most satisfactory approach is to assume that the chain crystallises with no gross distortion of the random coil which exists in the melt, though local changes must occur. A search for other models which fit the data is being made.

These results only apply to quenched samples as the clustering disqualifies results on slowly cooled and solution crystallised samples. Defenders of chain folding can thus say that it will still apply to crystallisation nearer to equilibrium. This is unsatisfactory so long as there is no prediction as to where 'near equilibrium' ends. One major omission of theories to date is that there has been no comprehensive attempt to treat the dynamics of the chain as it attaches to the crystal. If crystallisation is slow the chain should have plenty of time to relax its shape

to a 'jumping jack' as it is attached but when the crystals are growing rapidly they may well just entrap chains with little relaxation from the random coil. At the moment there is no basis for saying where the boundary between these two regimes lies. Chain folding theories assume that the first case applies over most of the normal range, my personal opinion is that the second case may be the rule, possibly even in solution crystallisation.

In the next year we may expect more neutron results on other systems. Ideally a polymer may be found where the labelled and unlabelled chains do not segregate significantly during slow crystallisation but since the chain motions involved in segregation and chain folding must be related we may reach an impasse. The adjacent re-entry chain folding model has received a nasty blow. If it does not survive as the description of melt crystallisation much of our understanding of polymer structure and properties will have to be reassessed. □

Radioastronomy and cosmology

from M. S. Longair

The cosmological problem as it presents itself to the radio astronomer was the subject of IAU Symposium No 74, Radio Astronomy and Cosmology, which was held at the Cavendish Laboratory, Cambridge on August 16–20, 1976.

THE standard textbooks on cosmology describe how the density, deceleration and geometry of the Universe may be determined from observations of the most distant galaxies. The remarkable fact that even the brightest extragalactic radio sources are at very great distances means that in principle they can be used in all the classical cosmological tests. In practice radio sources tell us rather little directly about these properties of the Universe because the overall properties of the radio source population evolve rapidly with cosmological epoch. Radio sources are associated with the most violently active systems— the nuclei of massive galaxies and quasars—and in general terms the distribution of radio sources traces how this violent activity has changed with cosmic time. All the evidence suggests that the Universe went through a period of violent activity when it was about one-tenth of its present age. Thus the contribution of discrete radio sources to cosmological problems is directed much more towards physical

aspects of the evolution of the Universe than towards the geometrical questions of classical cosmology.

The emphasis of the symposium was on the ways in which observations of discrete radio sources provide information of cosmological relevance. In interpreting the data, attention was focused on standard evolutionary world models.

The overwhelming impression was of the wealth of high quality observational data which has become available in recent years at radio and optical wavelengths. The radio observations span a range of 10^4 in sensitivity and the associated optical objects span a range in redshifts of at least 0 to 3.53. The radio surveys have been directly responsible for the discovery of large numbers of very large redshift quasars.

All such cosmological studies are ultimately based on surveys of radio sources and very large samples of sources are now available at frequencies of 0.408, 1.4, 2.7 and 5 GHz. Perhaps the simplest of all cosmological tests at radio wavelengths is the counts of radio sources which are shown in Fig 1 in differential form. The number of sources ΔN in the flux density interval ΔS is shown normalised to an arbitrary Euclidean count $\Delta N_0 \propto S^{-5/2} \Delta S$. A major advance has been the great increase in the numbers of sources counted at flux densities $S \leqslant 1$ Jy at all frequencies. The be-

haviour of the counts at 0.408, 1.4 and 2.7 GHz is similar, showing three distinct regions: at high flux densities, $S>3$ Jy, the observed numbers of sources increase relative to the Euclidean prediction with decreasing flux density; at intermediate flux densities, $0.1<S<1$ Jy, the numbers follow closely the Euclidean prediction; at the lowest flux densities, $S<0.1$ Jy, the numbers of sources converge relative to the Euclidean prediction. It has been known for many years that this behaviour contradicts the result expected for a uniform spatial distribution of sources which is that the counts should be a monotonically decreasing function of decreasing flux density and is evidence that the properties of the radio source population have evolved with cosmological epoch. As the frequency of observation increases, the "maximum" in the normalised differential counts broadens appreciably due to the presence of a larger proportion of sources with flat radio spectra. At 5 GHz this broadening is so marked that the counts follow closely the Eucliden prediction over a wide range of flux densities. In view of the small difference in frequency between 2.7 and 5 GHz, such a difference is remarkable and suggests that a new class of radio source with flat radio spectra appear at flux densities $S_5 < 0.1$ Jy. Several lines of evidence suggest that this new class may be weak radio galaxies.

In the past there have been suggestions that there are anisotropies in the distribution of sources over the sky. The large surveys of sources made by observers at Bologna, Cambridge, Molonglo, NRAO and Parkes have now been analysed using the powerful technique of power spectrum analysis and no evidence for anisotropies has been found. The Westerbork observers reported the results of many surveys made in the flux density range $S_{1.4}<0.1$ Jy where the source counts converge and no evidence of anisotropies was found. A few claims of anisotropies were made—for example, differences in the total numbers of sources detected in different sections of the Molonglo and NRAO surveys which are significant at about the 1% level. These results must however be compared with the large body of data which reveals no significant evidence of anisotropy. A. S. Webster (Institute of Astronomy, Cambridge) pointed out that this provides the best evidence for isotropy of the distribution of discrete objects in the Universe on scales ~ 1 Gpc.

The angular diameter–redshift relation is an attractive cosmological test because all the standard world models predict a minimum in the angular size of a standard rod with increasing redshift. Since radio sources extend to large redshifts, many observers have searched for such a minimum using the physical size of radio sources as a standard rod. R. C. Roeder (David Dunlap Observatory, Toronto) reminded observers that this result is only true if the matter distribution in the Universe is uniform; in completely inhomogeneous models, the minimum disappears. Observations of the relation for radio galaxies and quasars with redshifts up to 2.8 were reported by groups from Cambridge, NRAO and Westerbork. All that could be said at present was that the angular size distribution of quasars observed at large redshifts appears to agree with what is predicted from the known physical size distribution at small redshifts although the statistics are still small.

A slightly different result is found for the related test, the angular diameter—flux density relation. R. D. Ekers (Kapteyn Laboratory, Groningen) extended the earlier work of G. Swarup and his colleagues (Radio Astronomy Centre, Ootacamund). The median angular size of complete samples of radio sources decreases with decreasing flux density until the very lowest flux densities, $S \lesssim 0.1$ Jy, at which the median attains a roughly constant value. It is not yet known if this is a minimum or an asymptote. Calculations by V. Kapahi (Tata Institute, Bombay and Groningen) showed that even the decrease in the median angular size with decreasing flux density disagrees with the postulate that the physical sizes of radio sources are independent of redshift. His results suggest that the median physical size must be smaller at large redshifts, roughly by a factor $(1+z)^x$ where $x \approx$ 1 to 1.5. This result does not necessarily contradict the direct observations of the angular diameter–redshift relation for quasars—Kapahi's result refers to all radio sources, the other to quasars alone.

Optical identification of radio sources has proved to be one of the most successful techniques for finding very distant galaxies and quasars. R. Fanti (Radio Astronomy Laboratory, Bologna) and C. Perola (Institute of Physics, Milan) showed that for weak radio galaxies there is a correlation between their radio and optical luminosities but for powerful radio galaxies, as has been known for many years, the dispersion in absolute optical magnitudes is small and is independent of radio luminosity. This means that the galaxies associated with powerful radio sources may be used in the optical redshift–magnitude relation. H. E. Smith (Physics Department, La Jolla) showed that in the past few years this approach has led to the discovery of eight radio galaxies with measured redshifts in the range 0.467 to 0.752
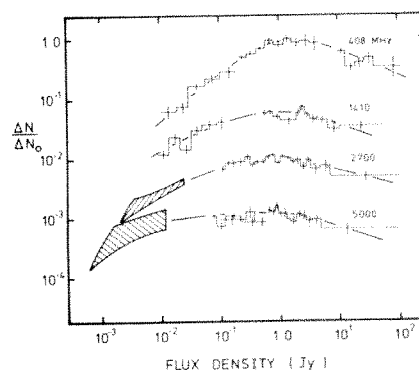
Fig. 1 The differential counts of radio sources at 0.408, 1.4, 2.7 and 5 GHz. The hatched areas are counts inferred from analyses of "confusion" surveys, that is, surveys in which the surface density of sources approaches one per beam area of the radio telescope. The ordinate scale is arbitrary.

and to the possibility of extending the redshift–magnitude relation for galaxies to significantly larger redshifts than before. He was not yet prepared to derive a value for the deceleration parameter from these data. For the quasars, J. N. Bahcall and E. L. Turner (Institute of Advanced Studies, Princeton) showed that despite the enormous scatter, their redshift–magnitude relation is consistent with the predictions of standard world models.

Given this increase in data, what can now be said about the evolution of the radio source population with cosmological epoch? M. Schmidt (California Institute of Technology) showed that for all available complete low frequency samples of quasars and radio quiet quasars the V/V_{max} test gives consistently values ≈ 0.65 as opposed to the value of 0.5 expected for a uniform distribution. The implication is that the quasars are found preferentially towards the limits of their observable volumes implying a strong evolutionary effect. One intriguing new result, as yet based on rather small statistics, is that the test gives the value 0.5 for flat spectrum quasars selected at high frequencies. If substantiated by improved statistics, this result will be difficult to reconcile with a simple evolutionary scheme because these quasars span a similar redshift range to those found in the low frequency surveys.

The counts of radio sources provide further constraints upon admissible evolutionary models. J. V. Wall (MRAO, Cambridge) showed that the improved statistics of the source counts make it much more difficult to find satisfactory models of the source distribution. All models require the strong evolution of only the most powerful classes of radio sources with cosmo-

logical epoch but it is not yet clear whether an abrupt cut-off to the distribution at large redshifts is necessary. Some of these questions may be answered by deep optical identification surveys of faint radio sources. The greatly increased data base provides new stimulus to the construction of models for the evolution of the radio source population.

There remain all the physical questions of the origin and evolution of radio sources and quasars. Many new facts emerged during the symposium: J. G. Bolton (CSIRO, Parkes) showed how three types of optical quasar could be distinguished on the basis of distinctive radio properties; A. Boksenberg (University College, London) showed the wealth of information contained in high resolution absorption line spectra of large redshift quasars and suggested the lines could be associated with intervening galaxies, a view opposed by E. M. Burbidge and H. E. Smith (Physics Department, La Jolla) who gave reasons for believing them to be intrinsic to the quasars; D. E. Osterbrock (Lick Observatory) described how detailed optical spectral studies of samples of radio galaxies enabled the properties of the emitting regions to be related to those found in other types of active galaxy.

The fundamental problems of understanding the origin and evolution of the radio source population and its significance for the origin and evolution of galaxies are far from being solved. It was clear from the symposium, however, that the impact of radioastronomy upon all aspects of cosmology is enormous. Detailed coordinated programmes of observation at radio and optical wavelengths of all types of radio source are now under way and this fruitful cross-fertilisation is one of the most encouraging prospects for the future of observational cosmology.

Infectivity marker for hepatitis

from Arie J. Zuckerman

INFECTION with hepatitis B virus is associated with the appearance in the serum of hepatitis B surface antigen (originally referred to as Australia antigen) and its homologous antibody. A second antigen, located in the core of the 42–45mm double-shelled spheroidal Dane particle, is intimately associated with the infection and gives rise to the core antibody. These markers have proved extremely valuable in unravelling the epidemiology of hepatitis B

and establishing the global dissemination and importance of this infection (see *Nature,* **247,** 2; 1974), have made possible the routine screening of blood donors for the surface antigen, and resulted in remarkable advances in knowledge of the virology and immunopathogenesis of hepatitis B and its associated chronic liver disease. In addition, the spread of infection can now be controlled in given circumstances, and more recently there have been far-reaching developments which include passive and active immunisation.

But despite the remarkably rapid application of these laboratory findings in clinical practice and preventive medicine a major problem remains unsolved. Hepatitis B virus has not been successfully propagated in tissue culture so that an *in vitro* assay system for infectivity is not available, and in its absence progress towards determining infectivity has been extremely limited. Only two species are known to be susceptible to hepatitis B, man and the chimpanzee. Infectivity cannot, and indeed must not, be experimentally measured in man, although some information can be deduced retrospectively in certain epidemiological settings and as a result of the application of newly developed sensitive serological markers which indicate exposure to the virus. Hepatitis B surface antigen and surface antibody were found in the serum of chimpanzees imported from Africa and living in captivity, suggesting that this species may be susceptible to hepatitis B infection. Experimental transmission of hepatitis B to a small number of previously unexposed chimpanzees established that the disease resembles clinically that in man and that the serological responses are identical. Although chimpanzees will be essential for assessing the infectivity, efficacy and safety of experimental hepatitis B vaccines now under development, there are many practical disadvantages to using them as animal models including their limited availability and expense. The world populations of the apes are already alarmingly small and these animals must be protected. Chimpanzees, although not as rare as the other apes, are being subjected to the combined threats of overhunting and the encroachment of civilisation, and they should be used with the greatest discretion for hepatitis studies (*Wld Hlth Org. tech. Rep. Ser.,* No 512, 1973; No 570, 1975).

Nevertheless, methods for determining infectivity are needed not only for the safety testing and evaluation of prospective hepatitis B vaccines but also for monitoring the removal of infectious virus from blood and blood products, for establishing physical and chemical methods of inactivation of

hepatitis B virus, and, of course, for the identification of infective carriers of the surface antigen of whom there are at a conservative estimate over 110 million (see *Nature,* **250,** 101; 1974).

It now seems that a potentially useful serological marker of relative infectivity may be available. This is a distinct soluble antigen, which although discovered in 1972 by Magnius and Espmark (*J. Immun.,* **109,** 1017; 1972), has until recently received little attention. The antigen—termed *e*—was found in some sera containing hepatitis B surface antigen and it differs immunologically and physicochemically from the previously described determinants of the surface antigen (Magnius, *Clin. exp. Immun.,* **20,** 209; 1975). Paradoxically, antibody against *e* has been found in serum specimens from healthy carriers.

The *e* antigen seems to be somehow intimately associated with the pathogenesis of liver damage. In other studies the *e* antigen was found by the immunodiffusion technique to be significantly more common in patients with chronic liver disease with persistent hepatitis B antigenaemia than in patients with acute viral hepatitis. Furthermore, the *e* antigen seemed to be a valuable prognostic marker since progression to chronic liver disease was recorded by serial liver biopsies in a number of patients with surface antigen-positive acute hepatitis associated with *e*. The clinical significance of the *e* antigen was supported by differences in the clinical, biochemical, and histological findings between patients with the *e* antigen and those without this antigen during the initial phase of viral hepatitis. The *e* antigen has also been found more frequently in surface antigen carriers with histological evidence of chronic liver damage. The presence of *e* antibody, however, is in general associated with only minor histological changes in the liver.

Other observations have also linked the *e* antigen with infectivity. It was commonly found in patients receiving treatment by maintenance haemodialysis and the antigen also appeared in the serum early in the incubation period of acute hepatitis B, at about the time of appearance of the surface antigen and before elevation of the serum transaminases. Furthermore, most of 12 healthy surface antigen carriers who had donated blood without producing clinical hepatitis in recipients possessed *e* antibody, which suggests that when this antibody is present the carriers may no longer be infectious.

Two very recent reports provide additional evidence for the relationship of the *e* antigen with infectivity. Okada and associates (*New Engl. J. Med.,* **294,**

Remarkable new beetle from central Italy

THE coleopterous fauna of Europe is undoubtedly better known than that of any comparable continental area in the world. By this time, to discover even a species new to science is a notable event in the life of a beetle collector; the discovery of a genuinely new genus in western Europe is something very few can hope to achieve. The finder of a new species representing a new taxon at higher than the generic level is almost unique in the last half century.

The fact that this has been achieved by Signor Roberto Pace of Verona, is not purely a matter of good fortune; he has been actively exploring a type of habitat which beetle collectors, except for a few in France, have largely neglected—the deeper layers of the soil. For the lack of an established English term for beetles of this type, we may adopt the French *endogées*. Endogean beetles have a good deal in common with cavernicolous ones—also studied mainly in France—and like them are scarcely represented in the British fauna. *Endogées* are usually small, eyeless, wingless, and more or less depigmented; from the obscurity of their habitat, very little is known about their foods or life cycles.

The insect discovered near Lazio in Central Italy by Pace (*Boll. Mus. civ. St. nat. Verona*, II, 445; 1975) represents not merely a new taxon of at least tribal level, but also a family and suborder hitherto unknown in the modern European fauna—and within the family and suborder represents a new and quite unexpected adaptive type. The suborder Archostemata, to which *Crowsoniella relicta* Pace belongs, is today a very small relict group, to which the oldest known fossils of Coleoptera (of Permian age, see Ponomarenko, (*The Historical Development of Archostematan Beetles*, Moscow, 1969)) have been attributed. Fossils of Archostemata are more or less numerous in "insect beds" throughout the Mesozoic era, but become much rarer in Tertiary deposits, more or less parallel to egg-laying mammals or Dipnoan lung-fishes.

Recent Archostemata are now placed in four small families (Crow-

son, *Boll Mus. civ. St. nat. Verona*, II, 459; 1975)· Ommadidae with one genus and two species in Australia; Tetraphaleridae with two species of *Tetraphalerus* in extra-tropical South America plus *Crowsoniella* in Italy; Cupedidae with four genera and about twenty species distributed through the main continental areas except for Europe and New Zealand; and Micromalthidae with a species in North America and a possibly distinct one in South Africa (Pringle, *Trans. R. ent. Soc. Lond.*, 87, 271; 1938). The larvae are definitely known to be woodborers in *Micromalthus* and a number of species of *Cupes* in America, Japan and Australia, and it is believed that this habit is general in the suborder. Except for *Micromalthus* and *Crowsoniella*, modern Archostemata are moderate-sized beetles with a body length rarely less than 7 mm, with well-developed wings, and active flight has been noted in several species.
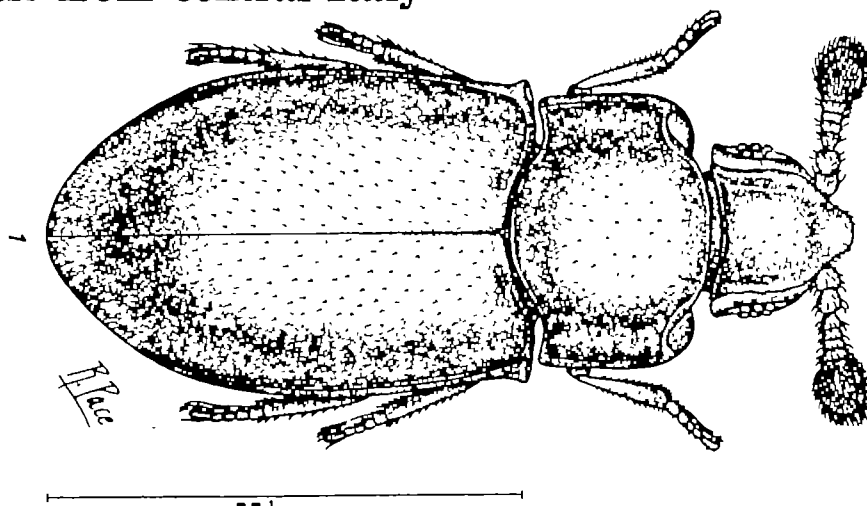
Apart from *Crowsoniella*, the only record of subterranean habits in the group is for *Micromalthus*, found breeding in timbers of more or less deep mines in South Africa (Pringle, *op. cit*). However, species of *Tetraphalerus* have been collected in Argentina (Monros and Monros, *An. Soc. cient. Argentina*, 154, 19; 1952) in areas of bushy vegetation with no large trees, which would seem to be unlikely to furnish the fungus-attacked dead trunks and logs which

are the usual breeding grounds of *Cupes* species; it seems quite possible that *Tetraphalerus* breeds in dead roots underground. If so, the habits of *Crowsoniella* might appear as a natural development from those of *Tetraphalerus*.

The antiquity of the *Tetraphalerus* line is indicated by well-preserved fossils, apparently close to the recent genus, in Jurassic deposits of the USSR (Ponomarenko, *op. cit.*), and the Tertiary occurrence of the type in Europe is shown by the fossil *Tetraphalerites oligocenius* (Crowson, *Ann. Mag. nat Hist*, (B) 5, 147; 1962) from the Bembridge limestone of the Isle of Wight. The *Crowsoniella* line may well have evolved its flightless and subterranean adults in the Mediterranean region during the Tertiary era. In having antennal grooves above rather than below the eyes, *Crowsoniella* differs from *Tetraphalerus* but may resemble some of the Mesozoic fossils described by Ponomarenko—the separation of the two lines may even be of Jurassic origin. Unfortunately the head is not preserved in the fossil *Tetraphalerites*.

Crowsoniella is surely one of the most ancient elements in the hypogean fauna of Europe; the possibility remains that related types will eventually be discovered elsewhere, e.g. in the Balkan peninsula, or even the Crimea or the Caucasus, if and when systematic search for hypogean forms is undertaken in those areas.

746; 1976) examined for *e* antigen and antibody, serum samples from twenty-three pregnant women who were asymptomatic carriers of hepatitis B surface antigen. Ten mothers with *e* antigen and seven with *e* antibody were identified. Their babies were tested at intervals for over a year after birth. All ten babies born to mothers with the *e* antigen acquired the surface antigen

which persisted throughout the period of observation, and interestingly enough all ten elder siblings of these newborn infants were also found to be asymptomatic carriers. In marked contrast, all seven babies born to mothers with *e* antibody escaped infection with hepatitis B virus and none of their three elder siblings were surface antigen carriers. It was concluded that the

e antigen may be used as an indicator of transmission of hepatitis B virus from carrier mothers to their children Interesting results were also obtained by Grady and colleagues (*Lancet*, ii, 492; 1976) during a survey of medical personnel accidentally inoculated with blood containing hepatitis B surface antigen. Hepatitis B infection or a surface antibody response occurred in

60% when *e* antigen was detected in the inoculum compared with 31% when *e* antigen was not detectable. The *e* antigen was detected in 74% of inocula accidentally sustained in maintenance renal dialysis and transplantation units, which are recognised hepatitis "high risk" areas, compared with 20% in other clinical units and the resulting incidence of hepatitis was 22% and 6% respectively. It was pointed out that if the correlation of *e* antigen and infectivity is indeed better than indicated by the results of this study, and the correlation is not merely an indirect reflection of some more specifically related phenomenon, then improvements in the sensitivity of the test will be of great value and the potential application to the problems of the transmission of hepatitis B virus will be almost limitless. □

Soil structure discussions in Adelaide

from I. J. Smalley

> A meeting of the Soil Physics Commission of the International Society of Soil Science was held at the Waite Research Institute, University of Adelaide, on August 23–27, 1976. The topic was Modification of Soil Structure and the proceedings edited by W. W. Emerson, R. D. Bond, and A. R. Dexter) will be published by John Wiley, London.

AUSTRALIA is the driest continent and *it is* also the saltiest. This fact causes great problems for Australian agriculture, and for the new town builders at Monarto, particularly as the irrigation water often carries a significant proportion of dissolved salt. But a salty clay can make a good cricket pitch, as J. R. Harris (CSIRO Division of Soils) pointed out in his paper on the maintenance of soil structure under playing turf. Different sports impose widely differing conditions. The soil structure of a cricket wicket requires fine structured clays of high bulk density to be maintained in a condition in which plasticity is retained when wet. When dry, the bounce height of the cricket ball and the ability to withstand intensive playing wear at the surface are important characteristics to match against soil properties. At the other extreme of the soil structure scale, bowling and putting greens are generally constructed on coarse-structured sands where the emphasis will be

on free drainage of surface waters and free rolling of the ball.

Discussion of sporting applications however, occupied only a very small proportion of the meeting; sessions were held on forces between colloidal particles, soil geometry, soil strength, soil tillage, soil management, soil permeability, measurement of aggregate stability, organic matter and aggregate stability, oxides and soil aggregation, organic material and soil structure, modification of soil structure with organic polymers—in all over fifty papers were presented.

The use of soil conditioners (admixtures to preserve a desirable soil structure) was discussed at some length and the meeting was to a large extent a continuation of the annual discussions on soil conditioners (see *News and Views*, **258**, 483; 1975). There is growing interest in conditioners as they become more economic and are seen to be effective. Soil conditioning is based on several more or less basic phenomena. One is aggregation of soil particles, which itself depends on adhesion between particles and binding material. In nature, the adhesive product is humus, an ill-defined polymeric material originated by organic matter. Soil conditioners are chemically similar and indeed provide adhesive forces between soil particles. J. P. Quirk and A. M. Posner (Universities of Adelaide and Western Australia) and their co-workers presented recent data on polyvinyl alcohol, particularly with respect to adsorption behaviour on different soil colloids. It seems that there are significant conformational differences between polyvinyl alcohol adsorbed on clay surfaces and on the hydrated aluminium oxide, gibbsite.

Some fundamental studies on soil structure and strength were reported. J. N. Israelachvili and G. E. Adams (Australian National University) have measured the forces between two mica surfaces in KNO_3 solutions. The results exhibit all the essential features characteristic of an interaction involving repulsive double-layer forces and attractive van der Waals forces (see also *Nature*, **262**, 774; 1976). R. C. Foster (CSIRO Division of Soils) presented ultramicromorphological data on some South Australian soils. Soils from irrigated pastures contained large amounts of organic materials, mainly plant remains in various stages of decomposition. Phytoliths were common, but the soil largely consisted of highly compressed and distorted cell wall fragments containing much polyphenolic material. Layers of clay particles down to 0.1 m thick were trapped between the organic layers. Near living roots and also sometimes at considerable distances from them, polysaccharide

mucigels up to 20 μm thick occurred.

J. E. Lloyd (University of Sydney) described a torsional shear box designed to measure the strength of seed beds. The identification of interest of participating engineers and agriculturists was here most noticeably achieved. In general the engineers at the meeting probably learnt more from the agricultural specialists than *vice versa*, but now that the engineers have decided that soil structure at the single particle level is important no doubt they will contribute more and more to this developing subject. □

A hundred years ago

THE obstruction at the entrance to New York Harbour known as Hell Gate was successfully removed by an explosion of dynamite on Sunday afternoon without any of the disasters that many people anticipated. The mass to be removed was about 70,000 cubic yards. The number of borings was 3,500; the number of galvanic batteries 200, placed in an explosion-proof chamber at a distance of 200 feet from Hell Gate. The diameter of the borings was uniformly 3 inches, and the depth varied according to circumstances, from 3 to 11 feet. Fifty thousand pounds of dynamite were used. The shock was not perceptible, not even glass being broken. A vast volume of water and smoke was driven about fifty feet into the air. All the charges were exploded, and the rock is stated to have been thoroughly removed. The explosion was heard at a distance of ten miles, and a tremor like a slight earthquake was heard in New York City and the localities contiguous to Hell Gate. The work has been in progress for seven years.

EARTHQUAKES were felt on the night of September 12–13, at Salonica, and in South Italy, at Reggio. Two motions were observed in the last city, the first one being the most notable, both having taken place on the 13th, between 12 and 1 o'clock, local time. Another earthquake was felt at Salonica, on the 14th, at 5 o'clock in the morning. The Reggio commotions were propagated to Messina and vicinity. They produced quite a sensation, although not destructive.

THE splendid orang-utang in the Berlin Aquarium died last week of consumption. Its friend and playfellow, the chimpanzee, died the next day of consumption and grief. The young gorilla, the one living specimen ever brought to Europe, which we referred to some months ago, is still alive, but ailing. Hamburg not long ago offered 100,000 marks for the gorilla; it is feared that he will soon be sold for less.
From *Nature*, **14**, September 28, 496, 497; 1876.

review article

Quantum field theory in curved space–time

P. C. W. Davies*

Recent theoretical developments indicate that the presence of gravity (curved space–time) can give rise to important new quantum effects, such as cosmological particle production and black-hole evaporation. These processes hint at intriguing new relations between quantum theory, thermodynamics and space–time structure and encourage the hope that a better understanding of a full quantum theory of gravity may emerge from this approach.

DURING the first quarter of this century, two major revolutions in physical science led to a complete restructuring of fundamental theoretical physics. The quantum theory met spectacular success in describing microscopic phenomena, such as atomic structure and chemical bonding. The theory of relativity, essentially a macroscopic theory, triumphantly explained the known discrepancy in the motion of the planet Mercury, and provided a basis for the understanding of cosmology. In its "special" form, applicable when gravity is negligible but velocities are large, the theory of relativity combines naturally with the quantum theory to yield a rich harvest of new results. The discovery of electron spin and antimatter arose directly out of this synthesis. So too did relativistic quantum field theory, which describes, for example, the interaction of electrically charged matter and electromagnetic radiation. This hybrid theory produced the modern conception of an elementary particle. The quantum field theory version of a particle differs in many respects from the classical notion of a little localised lump of energy. The prototype of the new version is the photon. It is the product of applying quantum theory to the electromagnetic field. Although the word "photon" conjures up the image of a tiny packet of light, this is quite misleading in a crucial respect. What is being quantised is not a localised object, but a mode of the wave field—something spread out over all of space.

The successful reconciliation of special relativity and quantum theory does not extend to the general case, where gravity is strong. Because the latter circumstances usually occur only in macroscopic systems, whereas quantum phenomena are microscopic, this mismatch has not been too serious in practice. In recent years, however, some interesting new possibilities have arisen, in which a proper understanding of quantum phenomena in strong gravitational fields might produce important advances in theory, and even lead to observational consequences.

The long term objective is to construct a full, consistent theory of quantum gravity, with the gravitational field itself quantised, perhaps after the fashion of the quantised electromagnetic field[1]. Even after many years of effort, this objective remains as elusive as ever. Sometimes it is conjectured that such theory would only be relevant on the minute length scale of 10^{-33} cm— the so-called Planck length, $(G\hbar/c^3)^{\frac{1}{2}}$—reducing the quantum gravity programme to a purely academic exercise. It now seems, however, that gravity may well lead to important quantum effects at much larger distances than this.

The basis for this new expectation is a study in which gravity

*Address: King's College, London University, Strand, London WC2R 2LS, UK.

is treated tentatively as a classical (unquantised) background field, in the presence of which all other fields are to be quantised. This somewhat easier approach is known as quantum field theory in curved space–time, because the unique status of the gravitational field as a geometric entity (the curvature of space–time) has encouraged a geometrical approach to quantum field theory, in the hope of a fresh perspective on formidable problems of principle. These problems always arise when the effects of external classical fields intrude into quantum field theory. It is hoped, though by no means certain, that the background field method represents a consistent first approximation, valid for distances much greater than the Planck length, to a (future) full theory of quantum gravity.

Although only tentative, this semi-classical treatment has nevertheless led to some unexpected discoveries, which may well point the way to a better understanding of the relationship between space–time structure and quantum theory.

Theoretical problems

Before describing those discoveries, it is necessary to explain the radical departures which the introduction of gravity into quantum field theory entails.

First, because the concept of "particle" is a global one, particle processes are sensitive to the global structure of the space–time. Ordinary quantum field theory depends heavily on the simple global properties of flat Minkowski space. Particles are defined successfully with reference to the underlying geometrical symmetries of this space. In a general curved space–time these symmetries are absent, and the notion of "particle" is without obvious meaning. Added to this is the problem that in general relativity, space–time may have new features such as horizons and singularities which must be taken into account in any satisfactory theory of quantum fields in curved space.

In certain simple cases, it may happen that the space–time possesses sufficient symmetry for a natural generalisation of the Minkowski space definition of a particle. This is so if the gravitational field is static, for example[2,3]. One distinctive state of a quantum field is that in which no particles at all are present. This is called the vacuum state. In Minkowski space the vacuum state is well defined, and corresponds intuitively to what we usually understand as a physical vacuum. In a more general space–time this is no longer true. What may be regarded as a vacuum in one reference frame (system of coordinates) may not be a vacuum in another frame. A very simple and dramatic illustration of this has been given by Stephen Fulling[3]. The essential features are present even in two-dimensional space–time, a diagram of which is shown in Fig. 1. An observer at

rest in the (t,x) frame follows the world line A. The other world line (B) corresponds to an observer who undergoes uniform proper acceleration to the right, for all time. His velocity relative to A approaches that of light for $t \to \pm \infty$. The light rays through the origin are drawn at 45° for convenience. Because information cannot propagate faster than light, these light rays act as an event horizon for B. He can never know about events which occur in the region marked "black hole". This appellation is used because the Universe in the direction of A would grow very dark as B zooms off to the right, on account of an escalating Doppler shift acting on any light travelling from A to B.

The space–time accessible to B is only a portion of the whole Minkowski space. It may be covered by a single coordinate system. In this portion, the space–time is still flat, but differs globally from Minkowski space because of the event horizon. It is a static universe, so that natural definitions of particle states and vacuum exist. Fulling showed that the vacuum state in this portion differs from the usual Minkowski space vacuum. This means that what A regards as a vacuum B does not. This has been interpreted by some[4] to mean that B will "see" particles, whereas A will not, a somewhat paradoxical idea. There is, however, no reason to suppose that the "particles" which B considers present correspond to "real" physical particles.
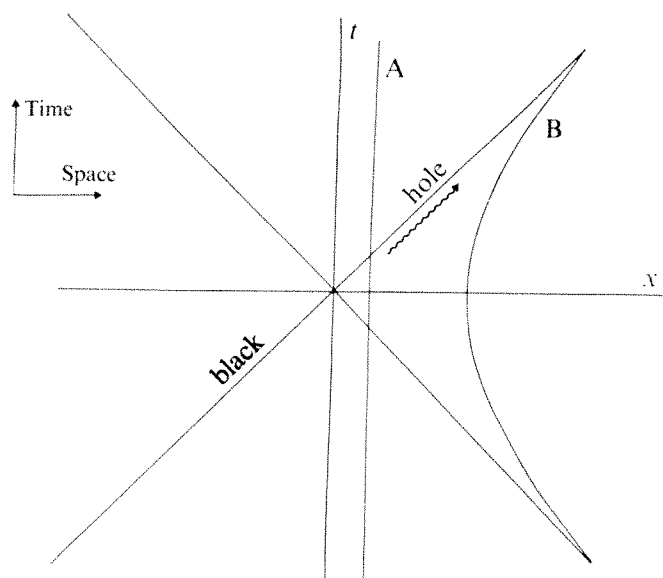
Fig. 1 The observer B undergoes uniform acceleration, while A remains at rest. The light ray through the origin to which B's trajectory is approaching forms an event horizon. B cannot see events above this line. This region of space–time is a "black hole"; light approaching B near this horizon becomes progressively more redshifted by the Doppler effect. The observer B has a completely different description of quantum particle states from that of A, who crosses into the black hole.

Using a simple geometrical argument, it is possible to show[5] that the spectrum of this mysterious radiation (in the massless case) is a Planck spectrum, corresponding to thermal equilibrium radiation at a temperature of $a/2\pi$, where a is the proper acceleration of the observer B. Recently, William Unruh has analysed the response of a simple model particle detector undergoing uniform acceleration, and confirmed that it does indeed detect thermal radiation at this temperature[6].

In this simple case, it is tempting to regard the unconventional vacuum as "unphysical". The enigmatic thermal radiation detected by B would then be dismissed as a disturbance caused by the agency which accelerates B. In a general space–time, however there is no easy way to single out one vacuum state over another. Nobody yet knows how to define a vacuum state in the general case.

The stress tensor

The ambiguities inherent in the concept of particles in regions of high space–time curvature has prompted an entirely different approach to quantum field theory in curved space[7]. In this new approach, one deals with mathematical objects which may be defined locally, and not globally as with particles. One important example is known as the stress–energy–momentum tensor, or stress tensor, for short. This object has the additional feature of acting as a source of gravity in Einstein's field equations, so that ultimately one might hope to be able to solve the theory of the coupled classical gravitational-quantum matter fields self-consistently.

Not all the difficulties associated with particle states are thereby circumvented, because the stress tensor must be calculated as an expectation value in some quantum state.

There are, however, several interesting situations in which, at some instant in the past, there exists a well defined vacuum state. The expectation value of the stress tensor can then be calculated in parts of the space–time which are strongly curved, as an expectation value in this initial vacuum state.

The simplest possible such case occurs if one starts with Minkowski space, and then disrupts it by introducing boundaries of some sort. This can be achieved in practice, for the electromagnetic field, using electrically conducting plates or mirrors, which reflect the field at their surfaces (see Fig. 2). If two such plates are held parallel and close, the Minkowski space vacuum is disturbed, even if the plates carry no electric charge. The reason is easy to understand. Modes of the field with wavelength greater than the separation of the plates cannot form as standing waves in the region between them, so they do not contribute to the vacuum energy.

It is easy to calculate the stress tensor between the plates. It was first done by Larry Ford[8]. The calculation does involve one fundamental and far-reaching subtlety, though. Formally, the stress tensor is infinite. This unpalatable circumstance is a result of the so-called "ultraviolet divergence", which is a universal problem in quantum field theory. It arises because even in the vacuum state, each mode of the field with frequency ω still has a zero-point quantum energy of $\hbar\omega$. Because the number of modes is infinite, the sum of all this energy is also infinite, progressively greater zero-point energy being contributed by the shorter wavelength modes.

In special relativity, this is no problem, because only differences in energy are observable. And the differences between empty Minkowski space and that with the parallel plates is finite. After subtracting away the divergent part Ford discovered that there was a uniform, static negative energy density between the plates. One consequence of this result is

Fig. 2 The Casimir Effect. Relative to the surrounding vacuum the space between the parallel conducting plates has negative energy, spread in a uniform cloud between the plates. It results in an electromagnetic force of attraction between the plates even though they are electrically neutral and there are no photons present.
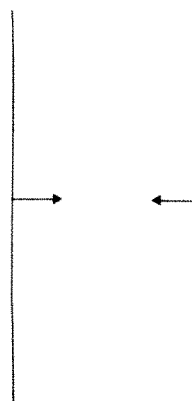
that an electromagnetic force of attraction should exist between the mirrors, even though they are electrically neutral. This effect had been known for many years, and is called the Casimir effect. It was predicted in 1948 by H. B. G. Casimir[9] and has since been measured experimentally[10].

If the mirrors are allowed to move about, a new phenomenon occurs. The static energy density between the plates is augmented by a flux of energy arising from the surface of the moving mirror. In a two-dimensional model, it is possible for the stress tensor to be calculated exactly in terms of functions implicitly dependent on the mirror trajectory[11]. If the system eventually becomes static once more, particle states may be defined in that future region, and the new phenomenon may be described as particle creation by the moving mirrors. This particle creation is present even for a single moving mirror.

The moving mirror models are useful in that they display in the simplest possible way some of the essential features of how the vacuum state may be disturbed by geometrical effects. When space–time curvature is allowed, similar phenomena occur.

Cosmology

Astronomy provides two scenarios in which gravity is strong enough to stimulate important quantum effects. The first is gravitational collapse, the second is big-bang cosmology.

The pioneering work on the cosmological case was carried out by Leonard Parker in the late 1960s. He showed that particles are created in an expanding universe, though exceedingly inefficiently in present conditions[12]. The motion of the universe acts in much the same way as the motion of mirrors, by disturbing the static vacuum. The more violent the motion, the greater the production rate.

During the early stages of the cosmological expansion, the expansion rate was enormously faster than now. At very early times, say 10^{-23} s—long before the Universe had even reached the reasonably well understood lepton era—this production may have been so prolific that it drastically modified the cosmological dynamics. In particular, if the expansion was anisotropic (different in different directions), then the effect of the particle production may have been to "smooth" out the anisotropy. It is a curious unexplained feature that the present condition of the Universe is one of high isotropy. There are hopes that this may now be understood as a result of the back-reaction of quantum effects. Very detailed studies[13,14] by the Russian astrophysicists Ya B. Zeld'ovich, A. A. Starobinskii and V. N. Lukas encourage these hopes.

It is important to realise that in the curved space case, the particle production occurs out of empty space. It does not arise from sources, or even from moving surfaces. The curvature of space–time itself may be caused by the presence of matter, but this matter does not directly cause the particle production.

Cosmological models have provided important "testing" scenarios for quantum stress tensor calculations (refs 13–16 and unpublished work of J. S. Dowker, R. Critchley, P.C.W.D. and S. A. Fulling). But now a host of new problems arises. Infinite terms which depend on the space–time curvature occur that cannot be removed by subtracting away the infinity which is already present in the Minkowski space case. Some of these terms have the same geometric structure as the left-hand side of Einstein's field equations, so they can be regarded as simply "renormalising" the values of the Newtonian gravitational constant G or Einstein's cosmological constant, albeit by an infinite amount. Other infinite terms cannot be so treated, and require a modification of the field equations to accommodate them. On top of all this the finite part of the difference (and perhaps some additional infinite terms) seems to contain ambiguities, which depend on the method by which the infinities are subtracted. There is no doubt that these ambiguities, which are probably inherent in the whole theory, present a major obstacle to understanding the subject. To remove the ambiguities, it is necessary to appeal to principles outside quantum field theory. Although there is no general agreement on what principles to invoke, very reasonable expressions for the stress tensor have been obtained by these methods (P.C.W.D., and S. A. Fulling, unpublished). For example, it seems certain that there is precisely no creation of photons in an homogeneous, isotropic universe, although there is probably still a non-zero vacuum energy density due to the space–time curvature.

Black holes

Gravitational collapse provides a rare opportunity for the application of quantum field theory because the space–time is approximately flat in the distant region where an observer would be located in practice. For this reason, the asymptotic situation may be described in terms of particles.

A study of this apparently very complicated problem has been carried out by Stephen Hawking[17], with very far-reaching consequences. The collapse, of a spherical star for example, disturbs the initial vacuum of the quantum field, and induces a flux of particles out of the collapsing object. The situation is closely analogous to the accelerated observer shown in Fig. 1. An event horizon forms around the collapsing matter; there is an infinite redshift from the surface. In this case B corresponds to the distant (inertial) observer who watches the collapse. The result is the same. He sees a flux of thermal radiation coming from the black hole. The temperature of this radiation is quite independent of the details of the star or its collapse history; it depends only on the mass M of the star through the formula

$$T = \frac{\hbar c^3}{8\pi GMk} \simeq 10^{-6} \left(\frac{M_\odot}{M} \right) K$$

More complicated expressions have also been derived for rotating and electrically charged black holes.

This astonishing discovery by Hawking has provided an enormous stimulus to the whole subject of quantum field theory in curved space. It also has profound implications in a wider sense. The created particles constitute a radiation flux in the asymptotic region where space–time is flat. This time the flux can, therefore, be unambiguously interpreted as a real flux of energy, coming out of the black hole. It follows that the black hole must shrink as its mass decreases. But as M falls, the temperature rises. Eventually the black hole apparently disappears in an explosive burst of radiation. In practice, only black holes of microscopic size, such as may have been formed early on in the big bang, will have suffered this fate, for a solar mass black hole has a temperature of a mere 10^{-6} K, and quantum effects are negligible.

The fact that black holes have a temperature implies that they are, in some sense, in thermal equilibrium. This means that they may be treated using thermodynamics. Direct analogues of the laws of thermodynamics were already known to apply to black holes even before Hawking's discovery[18]. Now these laws can be placed on a sound footing and new properties of black holes may be discovered[19]. For example, it turns out that rotating or charged black holes undergo a thermodynamic phase transition at high values of angular momentum or electric charge. Although it is only a conjecture, many people believe that Hawking's result heralds the discovery of important new connections between gravity, thermodynamics and quantum theory, with implications far beyond the subject of black holes.

Several enigmas remain. What does an observer who falls into the black hole see? How does the radiation react back on the hole, causing it to shrink? Where exactly is the radiation produced? What happens to the baryons which went to make up the collapsing star, which apparently disappears completely?

To answer these questions it is necessary to know what is happening near, and inside, the black hole, where the space–time curvature is comparable to the relevant radiation wavelengths and the particle concept breaks down. Some useful facts

emerge from a calculation[20,21] of the stress tensor. So far, this has only been done in two-dimensional models, though the essential features of the Hawking process are still present.

The results show that, in the reference frame of an observer at fixed distance from the centre of the star, the star seems to be surrounded by a cloud of static negative energy. This energy density is sometimes called vacuum polarisation, and is similar to that between the conducting plates in the Casimir effect. It is caused by the static space curvature around the star, and falls off rapidly with distance from its centre. If the star now implodes, a steady flux of energy is superimposed on this static cloud. At great distance this flux is the only contribution to the stress tensor and represents the Hawking radiation. The flux term arises mainly from the surface, and also the interior, of the star. This does not mean however, that particles are created there. Particles are not well defined in this region.

As the star shrinks inside the event horizon, the negative energy cloud around the star can stream into the black hole, causing its mass, and size, to decrease. The precise effect that this has on the collapsing matter is still uncertain.

The separation into flux and vacuum polarisation terms is really somewhat artificial, because it depends on the reference frame used by the observer. If the observer falls into the black hole, the situation is very different. He sweeps through the vacuum polarisation as he falls, so it appears to him no longer a static cloud but an outgoing flux of negative energy. By the time he reaches the event horizon, this flux is comparable in magnitude, but opposite in sign to the Hawking flux. Thus, he does not see much of the radiation which appears to his colleague out at infinity. This is consistent with the analogous situation shown in Fig. 1. There the inertial observer A who, you will recall, sees no particles in Minkowski space, corresponds to the falling observer who enters the black hole.

In practice, the quantum fields which are most important for the black-hole evaporation are the massless ones—electromagnetic, neutrino and, paradoxically, the gravitational field itself. If gravitons exist, they too will be radiated. Calculations[22] show that 2% of the luminosity will be due to gravitons, 17%

to photons and 81% to neutrinos. If the black hole rotates, the graviton flux is greatly enhanced. The Hawking effect thus produces the extraordinary situation that graviton processes are comparable in efficiency to electromagnetic processes. Rather than being restricted to the relatively uninteresting scale of 10^{-33} cm it is seen that on the much larger scale of say, 10^{-14} cm, full quantum gravity effects cause significant changes in the behaviour of strongly gravitating systems. These changes could be observed in principle; for example, by measuring the change in the orbital radius around the Sun of a mini black hole of mass about 10^{14} g. The rate of increase around the Sun of the orbital radius would depend sensitively on the mass loss rate, including that due to gravitons.

The new results of the past five years have enormously increased interest in quantum gravity. In spite of its highly esoteric nature, this subject is already seen to provide important new advances in hitherto unrelated subjects, such as black holes, thermodynamics and cosmology. The next five years can be expected to produce more advances still.

1 *Quantum Gravity—An Oxford Symposium* (edit. by Isham, C. J., Penrose, R., and Sciama, D. W.), (Oxford University Press, Oxford, 1975).
2 Ashtekar, A., and Magnon, A., *Proc. R. Soc.*, A346, 375 (1975).
3 Fulling, S. A., *Phys. Rev. D*, 7, 2850 (1973).
4 Gibbons, G. W., and Hawking, S. W., *Cosmological Event Horizons, Thermodynamics and Particle Creation*, Univ. Cambridg. preprint.
5 Davies, P. C. W., *J. Phys. A*, 8, 609 (1975).
6 Unruh, W. G., *Phys. Rev. D* (in the press).
7 De Witt, B. S., *Phys. Rep.*, 19, 295 (1976).
8 Ford, L. H., *Phys. Rev. D*, 11, 3370 (1975).
9 Casimir, H. B. G., *Proc. Kon. Ned. Akad. Wetenschap*, 51, 793 (1948).
10 Tabor, D., and Winterton, R. H. S., *Proc. R. Soc.*, A312, 435 (1969).
11 Fulling, S. A., and Davies, P. C. W., *Proc. R. Soc.*, A348, 393 (1976).
12 Parker, L., *Phys. Rev. Lett.*, 21, 562 (1968); *Phys Rev.*, 183, 1057 (1969); *Phys. Rev. D*, 3 346 (1971).
13 Zel'dovich, Ya. B., *Zh. Eksp. Teor. Fiz. Pis' ma Red.*, 12, 443 (1970) [*JETP Letts.*, 12, 307 (1970)]; with Starobinskii, A. A., *Zh. Eksp. Teor. Fiz*, 61, 2161 (1971) [*Sov. Phys. JETP*, 34, 1159 (1972)].
14 Lukash, V. N., and Starobinskii, A. A., *Zh. Eksp. Teor. Fiz.*, 66, 1515 (1974) [*Sov. Phys. JETP*, 39 742 (1975)].
15 Parker, L., and Fulling, S. A., *Phys. Rev. D*, 7 2357 (1973); 9, 341 (1974).
16 Streeruwitz, E., *Phys. Rev. D*, 11, 3378 (1975).
17 Hawking, S. W., *Comm. Math. Phys.*, 43, 199 (1975).
18 Carter, B., in *Black Holes* (edit. by DeWitt, B. S., and DeWitt, C.), (Gordon and Breach, 1973).
19 Davies, P. C. W., *Proc. R. Soc.* (in the press).
20 Davies, P. C. W., Fulling, S. A., and Unruh, W,G. *Phys. Rev. D*, 13, 2720 (1976).
21 Davies, P. C. W., *Proc. R. Soc.* A351, 129 (1976).
22 Page, D. N., *Phys. Rev. D*, 13, 198 (1976).

articles

Geology and late Cainozoic lake sediments of the Suguta Trough, Kenya

P. H. Truckle

Geology Department, Bedford College, Regent's Park, London NW1 4NS, UK

A consideration of the geological evolution and late Cainozoic lake deposits of the Suguta Trough in Kenya shows that for the later part of its history the growth of Lake Suguta was probably independent of that of Lake Rudolph, but was influenced instead by climatic change.

THE Suguta Trough (Fig. 1) is the northward continuation of the axial trough of the Kenya Rift. At this latitude the main rift is an ill-defined wide zone of faulting 130 km across. Northwards the trough passes into the southern end of the Lake Rudolph (renamed Lake Turkana) Basin,

and has its structural continuation in the Lake Stephanie Trough[1]. At present the Suguta and Lake Rudolph troughs are separated by the Barrier Volcanic Complex. The northern end of the Suguta Trough is an arid area of internal drainage occupied by the saline Lake Logipi and subject to seasonal flooding. The surface of this lake is 118 m below the 1968 level of Lake Rudolph[2]. The Suguta river is the only permanent supply of surface water to the trough, but Lake Logipi is maintained during the dry season by sub-surface flow through the alluvium flooring of the trough.

Only the upper part of the volcanic sequence is described here (Pliocene to Recent); it comprises alternating basalt
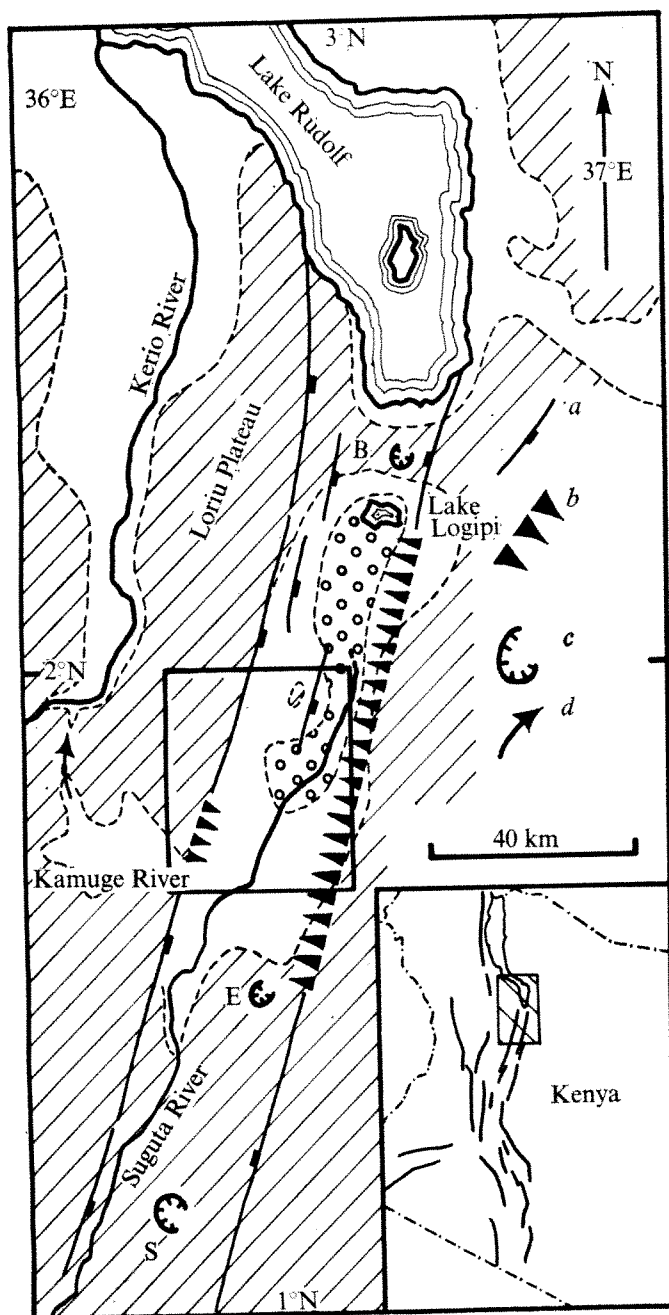
Fig. 1 The Suguta Trough. *a*, Normal fault; *b*, faulted monocline; *c*, caldera volcano; *d*, overflow of Lake Suguta; S, Silali; E, Emuruangogolak; B, the Barrier. Open circles, land under 305 m; diagonal rule, land over 606 m. Inset: geographical location of Fig. 1; centred box, Fig. 2.

and trachyte volcanism. The Suguta Trough was first initiated less than 3.0 Myr ago, and was subsequently filled in by younger volcanic products. The reforming of this trough, less than 0.75 Myr ago led to the eventual development of Lake Suguta.

Champion[3] has noted the occurrence of sediments flooring the trough and of wave-cut platforms on the southern side of the barrier, and has concluded that the two lakes were continuous before they became separated by growth of the barrier. Once separation was complete, Lake Suguta disappeared because of the insufficient water supply from the Suguta River relative to the evaporation rate. Reconnaissance studies of the Suguta sediments, and preliminary determinations of strand line altitudes, indicate that around 9,600 yr BP Lakes Sugata and Rudolph occupied apparently separate basins. The Recent history of Lake Suguta may have been dependent on climatic change rather than on an enlarged Lake Rudolph overflowing into the Suguta Trough.

Geological evolution

The geology of the Western Suguta Scarp is shown in Fig. 2.

Nathelot Basalts. These are the Lopiripira basalts of Rhemtulla[4]. A new name for this formation is proposed since the hill of Lopiripira is composed entirely of Nasaken trachyte agglomerate, and the basalts are typically developed at Nathelot. The basalts lie over older, tilted volcanic sequences, and the lava pile has the form of a large, low-angled shield. Nasaken trachytes underlie the Nathelot basalts and have yielded ages largely in the range 5.7–5.4 Myr, but younger dates for this volcano overlap with the age for the basal basalt flows[5]. A single basalt flow from this formation has given an age of 4.5 ± 0.4 Myr.

Loriu Trachyte Group. These trachytes lie conformably over the Nathelot Basalts and comprise three separate quartz trachyte volcanoes (Lopogut, Lomi and Kalakopon) whose alignment parallels that of the Suguta Trough. South of the Kamuge River, the Kapatagne Trachytes occupy a similar stratigraphic position[4]. The trachytes have given ages of 3.0 ± 0.2 Myr (Lomi), 4.4 ± 0.2 Myr (Kalakopon) and 3.3 ± 0.2 Myr (Kapatagni). The Tirr Tirr Trachyte and Basalt Group occurring on the eastern margins of the trough[6,7], has been dated at 3.6–3.9 Myr (ref. 8), suggesting it is broadly contemporaneous with the Loriu Trachyte Group. These trachytes are the northern and eastern extension of the South Turkana trachytes of Webb and Weaver[5].

Fig. 2 Geology of the Western Suguta Scarp. *a*, Nathelot Basalts; *b*, Loriu Trachytes; *c*, Lorikipi Basalts; *d*, Namarunu and Kamuge Trachytes; *e*, Kamuge Basalts; *f*, Suguta Basalts; *g*, Suguta Beds; *h*, member C basalts; *i*, member C pyroclastics; *j*, Recent basalts; *k*, boulder beds; *l*, alluvium.
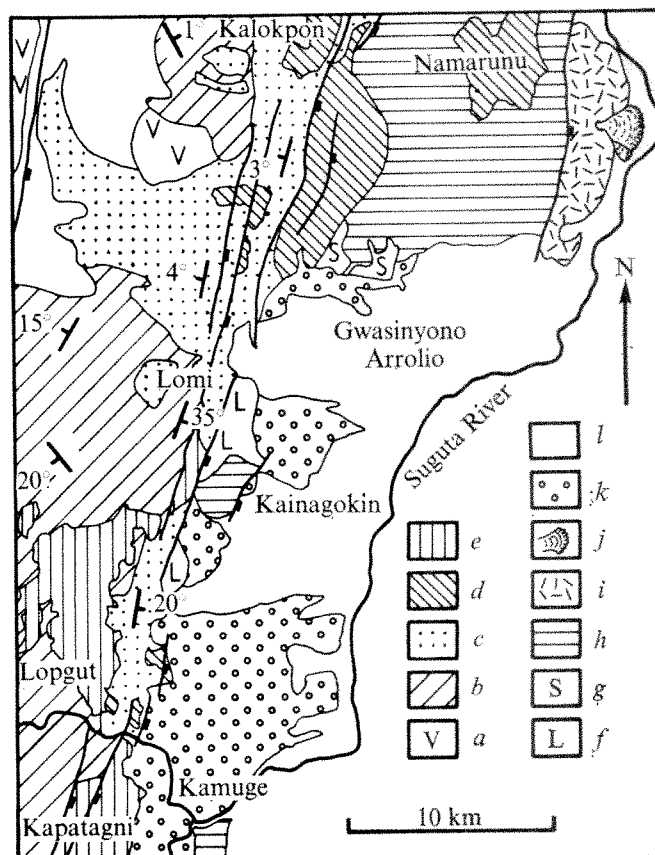
Monoclinal downwarping of the Loriu and Tirr Tirr groups led to the initiation of a proto-Suguta Trough.

Lorikipi Basalts. This basalt formation erupted from the floor of the trough, infilling and locally overflowing the structure. A single K–Ar determination of 4.0±0.1 Myr cannot be reconciled with the observed sequence.

Namarunu and Kamuge Trachytes. To the north, a mottled green trachyte flow and welded tuff sequence can be traced from the Loriu Plateau to the base of the Namarunu Trachyte Volcano some 300 m lower within the Suguta Trough. These trachytes lie over the Lorikipi Basalts but predate the next phase of faulting. To the south, the Kamuge Trachytes occupy a similar strati-graphic position, and a welded ash flow tuff has yielded an isotopic age of 0.75±0.1 Myr. The Kamuge Trachytes correlate with the lower flows from Emuruangogolak, whose activity commenced before the Suguta Trough reformed.

Kamuge Basalts. These crop out on the Loriu Plateau and are downwarped and downfaulted into the Suguta Trough. A K–Ar determination of '0' Myr indicates the youth of these basalts.

Monoclinal downwarping and minor faulting (between the Kamuge River and Lomi), passing into normal step faults, recreated the western Suguta scarp. The eastern Suguta Scarp was formed by a monoclinal downwarp and associated faulting[6,7].

Fig. 3 Stratigraphic column of Suguta Beds.

White diatomite

Conglomerate
Impure diatomite

Grits and silts

Conglomerate

Laminated silts

Grits

Conglomerate

Sub–aerial lava

Laminated silts

Shaley diatomites with concretions

Conglomerate
Silts with concretions

Diatomites with graded bedding

Shaley diatomites

Conglomerate and grits

Laminated silts and sands

Conglomerates grits and laminated silts

Base not seen

Following this last event, volcanism became restricted to the floor of the Suguta Trough and included the Suguta Basalts, member C basalts, and the continued eruption of Emuruangogolak and the Barrier, although the detailed geology of this complex is not well known. Incision of the Kamuge River system preceded the deposition of the Suguta Beds.

The Suguta Beds

Sediments occurring within the Suguta Trough and post-dating the latest phase of trough formation, have been designated the Suguta Beds. They comprise three sedimen-tary and one volcanic member (Fig. 3). The Suguta Beds are well exposed and crop out over wide areas of the Suguta Trough floor.

Member A is observed only in the vicinity of Kamuge, and is a dominantly clastic sequence of variable particle size. Rapid lateral facies changes, numerous minor unconformities and washouts indicate a fluviatile environ-ment. Member A is overlain with marked unconformity by member B diatomites.

Member B is well displayed at Kamuge and Gwasinyono Arrolio, where it is overlain by member C basalts. A similar sequence is seen at Kainagokin. At Kamuge, an essentially 'pure' diatomaceous sequence is overlain by silts, pebbly silts and silty diatomites, and at Gwasinyono Arrolio, a uniform sequence of bedded diatomites and intercalated pillow lavas occurs. Locally, grits underlie these diatomites.

Members A and B were tilted and downfaulted towards the axis of the trough and subject to erosion before the eruption of member C basalt.

Member C is represented by subaerial basalts resting unconformably on member B. At Gwasinyono Arrolio these basalts are intercalated with basal member D sediments and pass into the more extensive basalt field, capping the Namarunu trachyte volcano. Similar relation-ships obtain on the northern side of Emuruangogolak. To the east of Gwasinyono Arrolio, and at Kainagokin member C is faulted, although the relationship of the tectonic episode to member D sediments is not clear.

Member D at Gwasinyono Arrolio is a lower clastic sequence which overlies member C and is itself overlain by a thin sequence of well-bedded diatomites. The grain size of the clastic sediments decreases towards the trough axis. Thin diatomaceous deposits overlie member C at Kamuge and Kainagokin. Infrequent diatomites overlie member C basalts on Namarunu at an altitude of around 600 m. They contain gastropod remains, and are asso-ciated with algal limestones and wave cut basalt agglu-tinate cones. The gastropods have yielded a ^{14}C date of 9,660±210 yr b.p. (unpublished).

Members B and D record two separate phases of lacustrine sedimentation. During each phase Lake Suguta initially deepened but its development was terminated by desiccation which led to faunal extinction and was followed by subaerial erosion and volcanicity. Remains of *Croco-dylus, Lates* sp. and, rarely, gastropods (*Melanoides tuberculata*) and bivalves were found within members B and D. Rolled and badly worn bones were noted in the conglomerates of member A. The presence of elements of the Lake Rudolph/Nilotic fauna within members B and D suggests a hydrographic link between Lake Rudolph and Lake Suguta.

Lake levels

It is not possible at present to evaluate the effects of any deformation which may have affected strand lines within the trough. Within the Suguta Trough, high level lake benches occur at altitudes of up to 600 m. The 9,660-yr

date (member D) is of the same order as the 9,000±yr, climatically induced high stand of many East African lakes[*]. On the eastern Suguta Scarp, lake benches have been reported up to 550 m (refs 6 and 7) and up to 575 m on the southern side of the Barrier[10]. The age of these benches is not known. Butzer and Thurber reported[11] heights of 450–455 m for the highest deposits of the Kibish Formation (Omo area, Lake Rudolph), whose age range (27,444–5,450 yr BP) encompasses the value for member D. Although the chronology of the high level strand lines within the Suguta Trough is not well known it is possible to make some inferences about the relationship of Lake Suguta to Lake Rudolph, based on the assumption that the strand lines of the Suguta Trough have not been significantly deformed.

The differences in altitude between the Lake Suguta and Lake Rudolph strand lines suggest that from member D times at least, the Barrier complex was sufficiently well developed to prevent the two lakes coalescing. The present minimum altitude of the Barrier is 834 m (ref. 10), and it is questionable whether direct overflow of either lake across this complex could have taken place during member D times. An alternative outlet for Lake Suguta could be established by flooding the Kamuge River system, which would overflow into the Kerio river and subsequently into Lake Rudolph. Along the Kamuge river, infrequent diatomites occur at altitudes up to around 600 m. Overflow into the Kerio River would be possible should Lake Suguta levels reach an altitude of around 606 m. The sedimentary record indicates two major cycles in the development of Lake Suguta. This is more consistent with lake growth independ-

ent of Lake Rudolph rather than the progressive desiccation of an arm of Lake Rudolph, cut off by the growth of the Barrier.

Finally, it should be noted that it was apparently the present arid nature of the Suguta Trough which impressed Champion[1]. Under the prevailing climatic conditions inflow into the Suguta Trough is insufficient to maintain a large permanent lake, but during the 9,000-yr BP wet phase, Lake Baringo is likely to have supplied increased amounts of water to the Suguta Trough[8]. This increase of overflow from Baringo would further supplement the growth of Lake Suguta, already enlarged by climatic change throughout its drainage system. The rivers of the Suguta system can be expected, during this phase, to have had more permanent flow regimes than at present.

1 King, B. C., and Chapman, G. R., *Phil. Trans. R. Soc.*, A271, 185–208 (1972).
2 Cankwell, R. A., *Geogrl J.*, 137, 157–164 (1971)
3 Champion, A. M., *Geogrl J.*, 85, 323–41 (1935)
4 Rhemtulla, S., *Geogrl J.*, 136, 61–73 (1970).
5 Webb, P. K., and Weaver, S. D., *Bull. Volcanol.*, (in the press)
6 Baker, B. H., *Rep. geol. Surv. Kenya*, 53 (1963)
7 Dodson, R. G., *Rep. geol. Surv. Kenya* 60 (1963).
8 Baker, B. H., Williams, L. A. J., Miller, J. A., and Fitch, F. J., *Tectonophysics*, 11, 191–215 (1971).
9 Bishop, W. W., in *The Late Cenozoic Glacial Ages* (edit. by Turekian, K. K.), 493–527 (Yale University Press, 1971).
10 Baker, M. J., and Lovenbury, H. T., *Geogrl J.*, 137, 349–360 (1971).
11 Butzer, K. W., and Thurber, D. L., *Nature*, 222, 1138–1142 (1969).

Significance of major Proterozoic high grade linear belts in continental evolution

F. Bryan Davies
Department of Geology, City of London Polytechnic, Jewry Street, London EC3N 1ND, UK

Brian F. Windley
Department of Geology, University of Leicester, Leicester, UK

Major Proterozoic high grade linear belts are broadly conformable and sub-parallel to Archaean trends. Some are marginal to Archaean cratons, others intracratonic. All belong to a major set extending across the middle to late Proterozoic supercontinent. They are broadly contemporaneous with hairpin bends of polar wander curves, and some controlled igneous activity. Some are intracontinental structures developed in Proterozoic times when the deeper levels of continental masses underwent internal movement without loss of general coherence.

HIGH grade mobile belts may have formed in the overriding thickened plate on one side of continent–continent collision boundary[1] or they may have developed by reactivation of earlier Proterozoic or Archaean 'basement' rocks[2,3]. Data

relevant to the problem include the lack of offset of earlier oblique structures across several African belts[4], palaeomagnetic apparent polar wander curves suggesting that the bulk of Africa was an integral crustal unit throughout the Proterozoic[5,6] and Laurentia in the early Proterozoic (Hudsonian)[7,8] and divergent polar wander curves that indicate that the Grenville Belt formed at a plate convergent boundary[7,9] (this is inconsistent with the basement reactivation model[10] and with other palaeomagnetic data[11]). Palaeomagnetic evidence[1-13] suggests that there may have been a supercontinent throughout Proterozoic times (from 1,000 to 2,200 and possibly 2,700 Myr BP), but this has been challenged[14].

We produce here a plot of high grade Proterozoic mobile belts on the proposed supercontinent (Fig. 1), which, although problematic, is probably more relevant than the Permo-Triassic assembly of Pangaea (Gondwanaland rotated 180° between the Cambrian and the Permian[15]). We have not plotted the late Proterozoic Pan-African belts, because
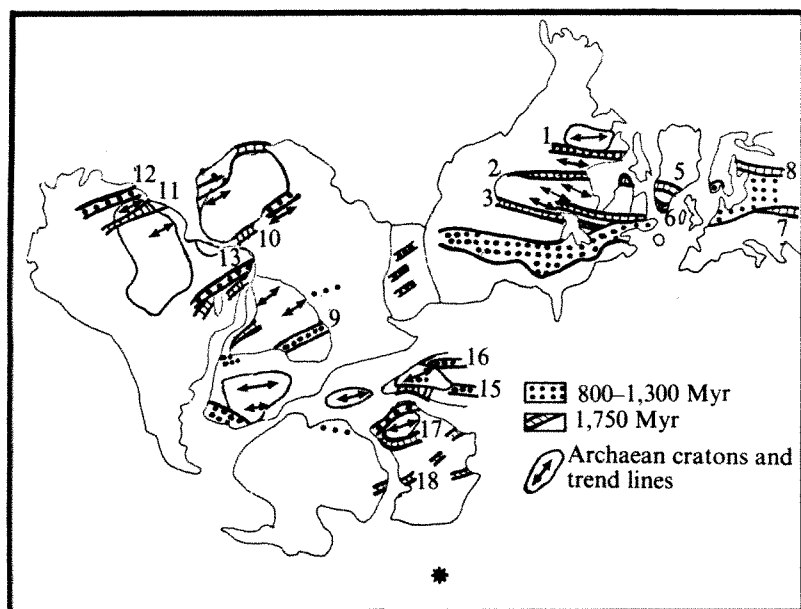
Fig. 1 Map of the mid–late Proterozoic supercontinent, palaeomagnetically determined by Piper[11–13], showing the distribution of two ages of Proterozoic lineaments lying on small circles about a point of rotation(∗) off the coast of Australia. Key to the major lineaments: 1, South-east margin of Slave Province[19,24]; 2, Nelson River, NW Superior Province; 3, SE Superior Province; 4, Grenville Front (and south-west extension)[10]; 5, Magssugtoqidian, Greenland (and unmarked Laxfordian of North-west Britain)[27–29]; 6, Ketilidian[27–29]; 7, Central Europe line[37]; 8, Karelide–Sveccofennide margin[38]; 9, Karagwe–Ankolan and Kibaran belts of central Africa; 10, Eburnian, east margin of West African craton[30]; 11, Amazon depression between Guyana and Brazillian Shield[35]; 12, North of Guyana Shield[35]; 13, Pre-Mines belts, Brazil[30,38]; 14, Eastern Ghats, east India[30,38]; 15, Satpura[30,38]; 16, Arivalli-Dehli belts, Peninsula India[39]; 17, Frazer belt[19,40]; 18, Adelaide belt[19,40].

so far they have been determined largely on radiometric evidence[16], nor the many low grade early Proterozoic geosynclines like Circum-Ungava and Hamersley, that contain banded iron formations[17] and probably formed as narrow Red Sea-type rifts according to palaeomagnetic[7,8] and stratigraphic evidence[18] with possible minor subduction. Neither of these types of belts, both discordant to our general trend, are considered here. In producing our plot (Fig. 1) we are neither attempting to validate the concept of a supercontinent nor denigrating the possibility that plate convergence contributed to the formation of some Proterozoic belts. We do suggest, however, that the sub-parallel distribution of so many belts in Fig. 1 places certain constraints on the movement of plates in the Proterozoic. Our principal aim is to emphasise that the high temperature nature of many of the mobile belts indicates that the continents suffered internal ductile strain under steep geothermal gradients often along transcurrent zones during the Proterozoic[19]. The lineaments we discuss could form either in the interior or margin of a thickened overriding plate of a collision-pair or within a large drifting continent or supercontinent[19]. Much structural[4,20] and palaeomagnetic evidence[5,6] favours the drift during much of Proterozoic time of several large continental sheets that suffered some marginal deformation giving rise, for example, to the Coronation[21] and Mount Isa[22] geosynclines, but more importantly, to internal ductile deformation along many discrete shear zones and mobile belts. Detailed structural studies[23] show that several major shear belts in western Greenland record transcurrent displacements that took place at some depth within a single continental mass. It is likely that this intraplate deformation is a reflection of the ductile nature of much of the Proterozoic continental mass —an intermediate stage between the widespread permobile activity in the Archaean and the more rigid behaviour of plates in the Phanerozoic, when mobile belts were confined to plate margins and internal deformation was brittle.

We use the term lineament here to include shear belts, mobile belts and linear zones of transcurrent displacements of magnetic and gravity anomaly patterns. These are high temperature belts, distinct from low grade geosynclines and high level faults.

Lineament development before 1,300 Myr BP

The lineaments of North America and the North Atlantic region are probably some of the best documented[19,24], and together fall into a number of types. Several intracratonic

lineaments strike north-eastwards across the Southern Superior Province and seem to be sub-parallel to the Grenville Belt. These lineaments deflect the Archaean structural, aeromagnetic and gravity patterns of the greenstone belts, and are traversed by undeformed members of the McKenzie Dyke swarm, dated at about 1,200 Myr BP. Some lineaments, particularly the craton marginal set were sites of granite emplacement or basic-acid volcanism during the mid-Proterozoic[25]. They therefore compare in broad age to the time span of such lineaments in the high grade terrains of the Canadian Shield as the western Superior, south-eastern Slave[26], and North Atlantic region. Of these, at least the Nagssugtoqidian of Greenland[23,27] and the Laxfordian of north-western Britain[28] were sites of transcurrent displacements and basic dyke emplacement in late Archaean and early Proterozoic times. Both these lineaments became the junctions between regions stabilised about 2,700 Myr BP and younger Proterozoic mobile belts[29]. In the later development of both lineaments, displacements perpendicular to them accompanied by granitic intrusions became prominent by about 1,600 Myr BP. It may be that this later stage is expressed in the Ketilidian Mobile Belt, which is dominated by displacements along the craton margin, and by considerable granite emplacement[27]. When plotted on a reconstruction of the continents before continental drift, the lineaments at the margin of the North Atlantic region, including those of eastern Scandinavia, fit the directions of both the intracratonic and marginal lineaments of the Canadian Shield[19,28]. Our argument is supported by the general coherence of palaeomagnetic poles from the Canadian Shield, which suggest that internal blocks did not move by large amounts relative to one another during the Hudsonian Orogeny when the shield seems to have undergone internal lineament deformation without loss of general coherence[7–9,14].

Early Proterozoic linear mobile belts are widespread in the Gondwana group of continents[11,19] and are mainly of the craton-marginal type. In Australia they form the junctions between the Yilgarn and Carpentaria blocks. Near-parallel structures occur in Antarctica, and the Eastern Ghats of eastern India. The Eburnean Belt lies along the eastern margin of the West African Craton and has a possible extension in South America, north of the Guyana Shield. The Ubendian and the older Limpopo belts in central and southern Africa are, however, oblique to this general trend.

The early Proterozoic lineaments seem to share a common evolution. Most probably they originated in the late

Archaean or early Proterozoic as transcurrent dislocations with associated basic dyke swarms which in several lineaments could fit the regional pattern. The development of the lineaments seems to have been arrested with the cratons at this early stage. Further development occurred in the lineaments along the margins of the cratons. In several lineaments of the North Atlantic Province a change from transcurrent displacements to movements perpendicular to the craton margins[37] was marked by an increase in granitic emplacement both along the lineaments and in broader zones parallel to them within the larger mobile belts. The culmination of mid-Proterozoic mobile belt development is probably marked by the intrusion of anorthosites, alkali complexes and rapakivi granites marking a major belt extending from the western USA to the Urals, which ceased not long after 1,300 Myr ago[35].

Lineament development between 1,300 and 1,000 Myr BP

Although the development of the last group of lineaments seems to have been arrested before 1,300 Myr BP, a number of narrow zones of basement reactivation and wider belts of tectonism and igneous activity developed between 1,300 and 1,000 Myr BP (ref. 30). The Grenville Belt and its possible extensions in North America and Europe is wider and more extensive than all other lineaments. Palaeomagnetic data[12] cast doubt on the concept that the formation of this belt involved some lateral plate motion[7-9].

A number of linear zones of this broad age can be outlined as anomalous structural, igneous and radiogenic features in Gondwanaland. These form a set of parallel lineaments concordant with the line of the contemporaneous Grenville Belt projected as a major zone extending through western Arabia, central Africa and along the eastern coast of South America (see Fig. 1), and with most of the earlier lineaments dating from around 1,700 Myr BP. Some belts, for example the Aravalli-Delhi Belt were active in both periods. The development of these 1,300–1,000 Myr lineaments has much in common with the later stage of development of the earlier lineaments, particularly in the association of major tectonism, the basic dyke emplacement about 1,200 Myr ago, and the syn- and post-orogenic granite emplacement and widespread thermal events.

Discussion

These Proterozoic lineaments seem to be coherent, global-scale, structures which belong to a single, broadly conformable pattern for about 1,000 Myr. Major events in the evolution of each age group of lineaments seem broadly contemporaneous with the hairpin bends of the single polar wander track for North America and Africa at 1,700 Myr BP and 1,300–1,000 Myr BP (refs 11 and 31), and therefore seem related to major changes in position of the Proterozoic continents or supercontinent. Many lineaments evolved as transcurrent dislocations within formerly continuous continental crust, rather than as sutures marking closed oceans.

In the context of a single Proterozoic supercontinent[11-13] both age groups of lineaments fit as zones of transcurrent displacement which lie on small circles about a point of rotation off the Australian part of the early continent. This agrees broadly with the interpretation of the Coronation[21] and Mount Isa[22] geosynclines and the Belt-Purcell Series of western North America as continental margins developed on the opposite leading edges of the rotated supercontinent.

Following the formation of the belt of anorthosites and rapakivi granites, the culmination of lineament development was expressed by late Proterozoic rifts, as in the Gardar Province of southern Greenland, which controlled lava extrusion, clastic continental sedimentation and the intrusion of alkaline complexes. These rifts may also be seen as the earliest expression of continental breakup connected with the eventual formation of the Proto-Atlantic Ocean in late Proterozoic times; they are comparable with the Triassic rifts containing continental sediments and volcanics along the Atlantic coast of the USA[32]. In fact, the large number of failed rifts and aulocogens of late Proterozoic age[33] is probably a demonstration of how difficult it was to fragment the Proterozoic supercontinent. Both types of Proterozoic lineaments are, on a global scale, sub-parallel to the trends of the Archaean granite–greenstone belts and therefore could be influenced by the overall Archaean structural pattern[34]. A few Proterozoic lineaments seem to have incidentally influenced the form of some Phanerozoic orogenic belts (for example, the old Amazon depression shield lineament segmenting the Andes along the Amotape Zone[35]).

Besides simple collision, there are several other means of explaining the formation of different Proterozoic high grade belts. Evidence from the shear belts is consistent with, but does not necessarily suggest, the idea that they are the deep-seated equivalents of modern intracontinental transform faults[36]. The Grenville Belt has been considered as the result of the slow drift of a continental plate over one or a series of hot spots, giving rise to successive stages of extension and compression. At deep levels of erosion an aulocogen may appear as a high grade ensialic mobile belt with some degree of transcurrent movement[42]. Upwelling and diverging convection currents may give rise to high heat flow and the formation of ensialic belts[43].

1 Dewey, J. F., and Burke, K. C. A., *J. Geol.*, **81**, 683–692 (1973).
2 Watson, J. V., *Phil. Trans. R. Soc.*, **A273**, 443–455 (1973).
3 Kroner, A., *Phil. Trans. R. Soc.*, **A280**, 541–554 (1976).
4 Shackleton, R. M., in *Implications of Continental Drift to the Earth Sciences*, 2 (edit. by Tarling, D. H., and Runcorn, S. K.), 1091–1095 (Academic, London, 1973).
5 Piper, J. D. A., in *Implications of Continental Drift to the Earth Sciences*, 1 (edit. by Tarling, D. H., and Runcorn, S. K.), 19–31 (Academic, London, 1973).
6 Piper, J. D. A., Briden, J. C., and Lomax, K., *Nature*, **245**, 244–248 (1973).
7 Irving, E., and Lapointe, P. L., *Geoscience*, **2**, 90–98 (1975).
8 Irving, E., and McGlynn, J. C., *Phil. Trans. R. Soc.*, **A280**, 433–468 (1976).
9 Irving, E., Emslie, R. F., and Ueno, H., *J. geophys. Res.*, **79**, 5491–5502 (1974).
10 Wynne-Edwards, H. R., in *Variations in Tectonic Styles in Canada* (edit. by Price, R. A., and Douglas, R. J. W.), *Spec. Pap. geol. Ass. Can.*, **11**, 263–334 (1972).
11 Piper, J. D. A., *Nature*, **251**, 381–384 (1974).
12 Piper, J. D. A., *Nature*, **256**, 519–520 (1975).
13 Piper, J. D. A., *Phil. Trans. R. Soc.*, **A280**, 469–490 (1976).
14 McGlynn, J. C., Irving, E., Bell, K., and Pullajah, G., *Nature*, **255**, 318–319 (1975).
15 Briden, J. C., Drewry, G. E., and Smith, A. G., *J. Geol.*, **82**, 555–574 (1974).
16 Hepworth, J. V., *Phil. Trans. R. Soc.*, **A280**, 662–663 (1976).
17 Goodwin, A. M., in *Implications of Continental Drift to the Earth Sciences*, 2, (edit. by Tarling, D. H., and Runcorn, S. K.), 1047–1070 (Academic, London, 1973).
18 Trendall, A. F., in *Genesis of Precambrian iron and manganese deposits, Proc. Kiev Symp.*, 1970 (Unesco, Paris, 1973).
19 Sutton, J., and Watson, J. V., *Nature*, **247**, 433–435 (1974).
20 Shackleton, R. M., *Phil. Trans. R. Soc.*, **A280**, 491–498 (1976).
21 Hoffman, P., *Phil. Trans. R. Soc.*, **A273**, 547–581 (1973).
22 Dunnet, D., *Phil. Trans. R. Soc.*, **A280**, 641–645 (1976).
23 Bak, J., Sørenen, K., Grocott, J., Korstgard, J. A., Nash, D., and Watterson, J., *Nature*, **254**, 566–569 (1975).
24 *Geological Survey of Canada*, 1 : 5,000,000 Aeromagnetic map (1968).
25 Bridgwater, D., and Windley, B. F., *Phil. Trans. R. Soc.*, **A273**, 493–512 (1973).
26 Douglas, R. J. W., and Price, R. A., *Spec. Pap. geol. Ass. Can.*, **11**, 626–688 (1972).
27 Escher, A., Escher, J. C., and Watterson, J., *Can. J. Earth Sci.*, **12**, 158–173 (1975).
28 Davies, F. B., *Nature*, **256**, 568–570 (1975).
29 Bridgwater, D., Watson, J. V., and Windley, B. F., *Phil. Trans. R. Soc.*, **A273**, 493–512 (1973).
30 Hurley, P. M., and Rand, J. R., *Science*, **164**, 1229–1242 (1969).
31 Donaldson, J. A., McGlynn, J. C., Irving, E., and Park, J. K., in *Implications of Continental Drift to the Earth Sciences*, 1 (edit. by Tarling, D. H., and Runcorn, S. K.), (Academic, London, 1973).
32 Klein, G. de V., *Bull. geol. Soc. Am.*, **80**, 1825–1832 (1969).
33 Burke, K., and Dewey, J. F., *J. Geol.*, **81**, 406–433 (1973).
34 Engel, A. E. J., and Kelm, D. L., *Bull. geol. Soc. Am.*, **83**, 2325–2340 (1972).
35 Sillitoe, R. H., *Nature*, **250**, 242–245 (1974).
36 Molnar, P., and Tapponier, P., *Science*, **189**, 419–426 (1975).
37 Windley, B. F., *Phil. Trans. R. Soc.*, **A273**, 321–341 (1973).
38 Watson, J. V., *Abstr. European geol. Soc.*, Reading (1975).
39 Read, H. H., and Watson, J. V., *Introduction to Geology*, 2 (Macmillan, London, 1975).
40 Crawford, A. R., *J. geol. Soc. India*, **10**, 117 (1969).
41 Baer, A. J., Emslie, R. F., Irving, E., and Tanner, J. G., *Geoscience*, **1**, 54–61 (1974).
42 Badham, J. P. N., *Phil. Trans. R. Soc.*, **A280**, 661–662 (1976).
43 Shackleton, R. M., in *Time and Place in Orogeny, Spec. Publ. geol. Soc. Lond.*, **3**, 1–8 (1969).

Sex hormone receptors in mammary tumours of GR mice

Mels Sluyser, S. G. Evers & C. C. J. De Goeij

Division of Endocrinology, The Netherlands Cancer Institute, Antoni van Leeuwenhoek-Huis, Amsterdam, The Netherlands

Studies of the growth behaviour and hormone receptor contents of GR-mouse mammary tumours suggest that the hormone responsive tumours are mixed populations of hormone-dependent cells and autonomous cells. The hormone-dependent moiety of these tumours contains high levels of progesterone receptor and oestrogen receptor. The autonomous moiety has a low but probably significant oestrogen receptor level, and is practically devoid of progesterone receptor. Androgen receptor levels in both moieties are low. Endocrine ablation prevents growth of the hormone-dependent moiety of the tumours, but not of the autonomous moiety.

IN approximately one-third of patients with metastatic breast cancer, the tumour responds favourably to ablative or additive forms of endocrine therapy[1]. Several groups of investigators have shown that a relationship exists between the presence of specific oestrogen-binding proteins in human breast cancer tissues and hormone responsiveness of the tumour[1-4]. But although assays of oestrogen receptors have important prognostic value, a large

proportion of the oestrogen receptor-positive group does ont respond to endocrine treatment[5]. Similarly, studies on DMBA-induced mammary tumours in rats suggest that determination of cytoplasmic oestrogen-binding capacity alone does not perfectly predict whether the growth of a tumour will be affected by ovariectomy[6].

It has been proposed that in non-responding receptor-positive cases the oestrogen receptor can bind hormone but cannot subsequently cause gene activation, and that assays of progesterone receptor may have greater prognostic significance[7]. It has also been suggested that determination of androgen receptor may be an important tool for predicting response to castration in premenopausal patients[8].

To evaluate the roles of different hormone receptors in mammary tumour growth, it is important to have available model systems with which the influence of each factor can be established. The GR-mouse system seems to be useful in this respect. Hormone-dependent mammary tumours can be induced in ovariectomised GR mice by continuous treatment for 3–4 months with oestrone and progesterone. The tumours can be transplanted into castrated $(O20 \times GR)F_1$ hybrid mice which are treated with oestrone and progesterone. The tumours generally retain their hormone dependence during the first transplant

Fig. 1 Serial transplantation of mammary tumours in castrated GR mice (— — —), and in castrated GR mice which were treated continuously with oestrone and progesterone (———). M is the tumour weight (mg) and t the time (d) after grafting Receptor contents of the tumours are expressed in fmol per mg cytosol protein. Receptors of progesterone (PR), oestrogen (OR) and androgen (AR).

Fig. 2 Serial transplantation of mammary tumours in castrated GR mice (— — —) and in castrated GR mice treated with oestrone and progesterone (———). M is the tumour weight (mg) and t the time (d) after grafting. Receptor contents of the tumours are expressed in fmol per mg cytosol protein. Receptors of progesterone (PR), oestrogen (OR) and androgen (AR).
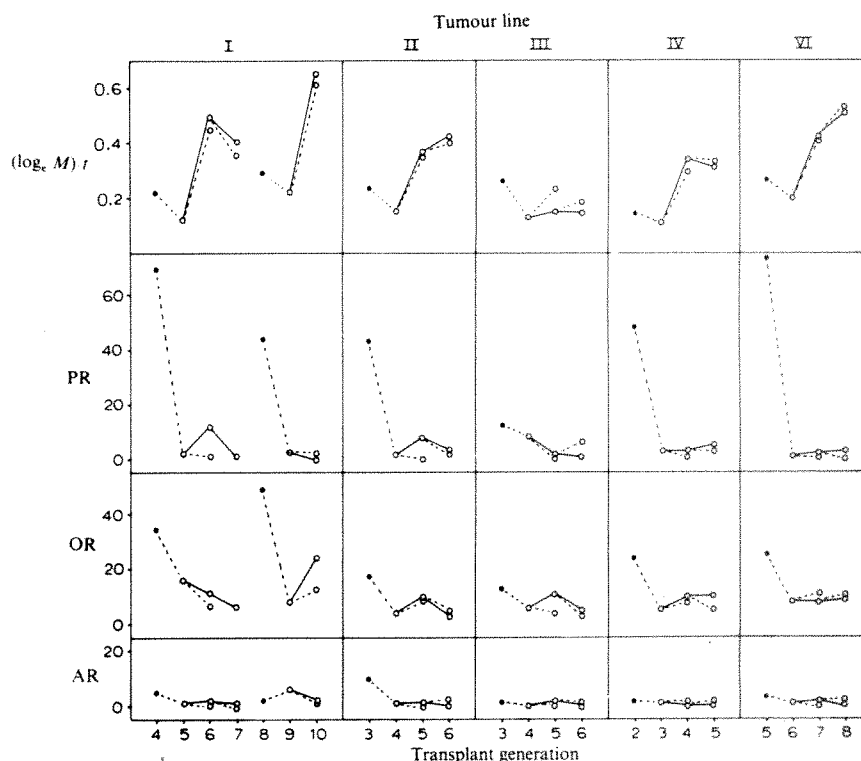
generation, but in subsequent transplant generations a progressive decrease in hormone dependence is observed which eventually leads to complete hormone independence. An advantage of using this system is that a certain tumour can be studied in various endocrine conditions. We have reported[9] previously that the mammary tumours which respond to oestrone and progesterone contain higher levels of oestrogen receptor than the autonomous mammary tumours. We now present results of a study on the relative concentration of oestrogen receptor (OR), progesterone receptor (PR) and androgen receptor (AR) in these tumours, and the relationship with tumour growth behaviour.

Transplantation series

Mammary tumours were induced in ovariectomised GR mice and tumour lines developed by serial transplantation from seven separate primary tumours (tumour lines I–VII). The degree of hormone responsiveness of the primary tumours and of each transplant generation was tested by grafting portions of the tumours in two castrated mice which were treated with oestrone and progesterone (OPC mice), and two castrated mice which were not treated with hormones (C mice). Tumours which were transplantable in OPC mice but which did not yield visible outgrowths within 3 months in the C mice were designated hormone dependent (HD). Tumours which were transplantable in both cases but which yielded outgrowths in OPC mice sooner than in C mice, were designated hormone responsive (HR). Tumours which grew equally well in OPC as in C mice, were designated autonomous or hormone independent (HI).

Samples of 500-mg tumour tissue were used for hormone-receptor assay. 2,4,6,7-n-^{3}H-Oestradiol (85 Ci mmol^{-1}) was obtained from the Radiochemical Centre. R5020 (17,21-dimethyl-19-nor-pregna-4,9-diene-3,20-dione), ^{3}H-R5020 (51.4 Ci mmol^{-1}), R1881 (methyltrienolone) and R1881-6,7-[3]H (58.2 Ci mmol^{-1}) were gifts of Dr J. P. Raynaud, Roussel-Uclaf. Nafoxidine hydrochloride (U 11,100A) was a gift from the Upjohn Company. Determination of oestrogen receptor in the tumour cytosols was carried out with the dextran-coated charcoal assay[10,11]. The same method using ^{3}H-R5020 was used for progesterone receptor assay[12], and with ^{3}H-R1881 for androgen receptor assay[13]. Controls contained excess non-radioactive ligand. Protein was assayed according to Lowry *et al.*[14]. Recep-

tor levels were expressed as fmol mg^{-1} cytosol protein. Tumour tissues stored at $-70\ ^{\circ}$C were used for receptor assays, since in cytosol preparations sometimes losses of OR and PR were found to occur up to 0.2 and 0.8% d^{-1}, respectively, when these were stored at $-70\ ^{\circ}$C.

Figure 1 shows the results obtained with transplantation series I–VII. As a parameter of the growth of each individual tumour we calculated the ratio $(\log_e M)/t$, where M was the weight (mg) of the tumour when it was collected (usually about 3,000 mg), and t was the time (d) after grafting. Six of the primary tumours were found to be HD, since they were transplantable in OPC mice but not in C mice; transplantation series I–VI were developed from these tumours. The seventh primary tumour was found to be autonomous, since its first transplant generation appeared equally rapidly in OPC mice and in C mice; this line (tumour line VII) was therefore discontinued. Tumour lines I–IV and VI exhibited a decrease in hormone dependency after the second to third transplantations, but line V remained HD up to the ninth transplant generation.

When HR tumours were grafted in C mice (Fig. 1, dotted lines), they gave rise to tumours with markedly lower $(\log_e M)/t$ values than when they were grafted in OPC mice (Fig. 1, solid lines). This was due to the fact that in the C mice the tumours appeared after a longer lag period than in the OPC mice. The tumours which appeared in C mice had significantly lower PR and OR contents than those which were obtained in OPC mice. Evidence that the tumours derived from C mice were autonomous was obtained by transplanting portions of 6 of these tumours in OPC mice and C mice. As is shown in Fig. 2, oestrone and progesterone did not cause an increase in the $(\log_e M)/t$ values of the transplants. Furthermore, the tumours remained HI during subsequent serial transplantations, whether or not hormones were given, and their receptor contents remained low.

These results indicate that in the absence of oestrone and progesterone, the HD cells in a hormone-responsive graft did not multiply, and that the HI cells in the graft consequently gave rise to autonomous outgrowths. The long lag period can be attributed to the long time required for the small percentage of HI cells in the graft to form a visible tumour (Fig. 3b). This suggests that all tumours indicated in Figs 1 and 2 by open circles are autonomous.

Hormone dependent   Hormone responsive   Hormone independent
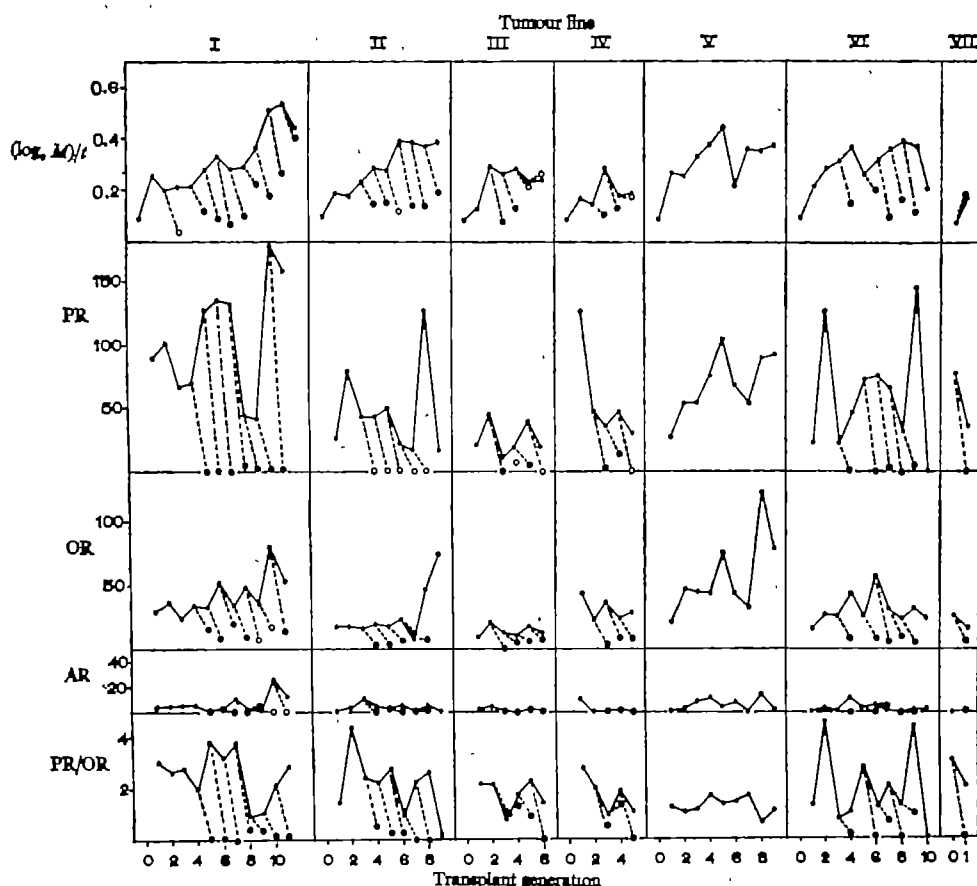(Autonomous)

Fig. 3 Serial transplantation of mammary tumours in castrated GR mice (C), and in castrated GR mice which were treated continuously with oestrone and progesterone (OPC). Tumours indicated by R were found to have high hormone receptor contents, those without R had low hormone receptor contents. A possible interpretation is given in the lower part of the figure. Shaded cells are hormone dependent, unshaded cells are autonomous. The scheme proposes that mouse mammary tumours progressively become less hormone dependent during transplantation because there is an increase in the percentage of autonomous cells in the tumours. When a hormone-responsive tumour is growing in

the presence of hormones (*a*) both the hormone-dependent cells and the autonomous cells can multiply, whereas in the absence of hormones (*b*) only the autonomous cells can multiply. Case *b* therefore yields an autonomous tumour with low receptor values. It is not clear whether the primary hormone-dependent tumour from which the transplantation series is derived is entirely devoid of autonomous cells or already contains minute amounts of autonomous cells. Both alternatives are illustrated at the bottom left.

Receptor levels and ratios

Figure 4 shows the average values and standard deviations of receptor contents of hormone-dependent, hormone-responsive and autonomous tumours. The levels of PR and OR were both markedly lower in HI tumours than in HD or HR tumours. No significant difference in hormone-receptor content was found between HD and HR tumours. The results indicate that PR levels below 17 fmol per mg cytosol protein and OR levels below 14 fmol per mg cytosol protein were markers for hormone independence. Androgen-receptor levels in HD and HR tumours were markedly lower than PR and OR levels in these tumours. AR levels in HD and HR tumours were on average slightly higher than in HI tumours. Figure 4 also shows the results obtained when the PR/OR ratio of each tumour was calculated separately, and the average values and standard deviations of the data were computed. HD and HR tumours had significantly higher PR/OR ratios than HI tumours. No significant difference in PR/OR ratios were found between HD and HR tumours. The results indicate that PR/OR values smaller than 0.8 were markers for hormone independence.

Discussion

Calculation of PR/OR ratios served as an additional check for the significance of absolute PR and OR values. A suitable marker for hormone independency in GR-mouse mammary tumours seemed to be PR and OR contents lower than 17 and 14 fmol per mg cytosol protein, respectively, together with a PR/OR ratio lower than 0.8. It should be pointed out that these values apply to the method used for receptor assay in which only the number of free receptor sites was determined. In the OPC

mice the endogenous levels of both progesterone and oestrogen would be very high, so the quantitative levels of free cytosol receptors could be small compared with the bound receptor.

At present it is not clear whether the level of cytoplasmic hormone receptor can be taken as a measure of the number of cells in the tumour that contains the receptor. To substantiate this conclusion, other methods such as autoradiography should be utilised. If this assumption is correct, however, our data are consistent with the assumption that hormone-responsive mouse mammary tumours are heterogeneous and that they contain the following cell populations. (1) Hormone-dependent cells with average PR, OR and AR contents of 74, 48 and 6 fmol mg^{-1} cytosol protein, respectively, and an average PR/OR ratio = 2.0. These cells depend for growth on oestrogen and progesterone. The hormone-dependent population might consist either of one cell type with all three receptors, or might contain different cell types with different receptor contents. There is evidence for the presence of different stem lines in hormone-dependent rat mammary tumours[15]. (2) Hormone-independent (autonomous) tumour cells with average PR, OR and AR contents of 3, 9 and 1 fmol mg^{-1} cytosol protein, respectively, and an average PR/OR ratio = 0.4. These cells are not affected by endocrine ablation.

The results obtained with the GR-mouse model system reveal some of the limitations of hormone-receptor assay as a parameter on which to predict whether the growth of a mammary tumour will be affected by endocrine treatment. One of these limitations is that whereas low PR, OR and PR/OR values seemed to be a good indication for hormone independence, high PR, OR and PR/OR values did not reveal whether or not the tumours contained autonomous cells in sufficient quantity to
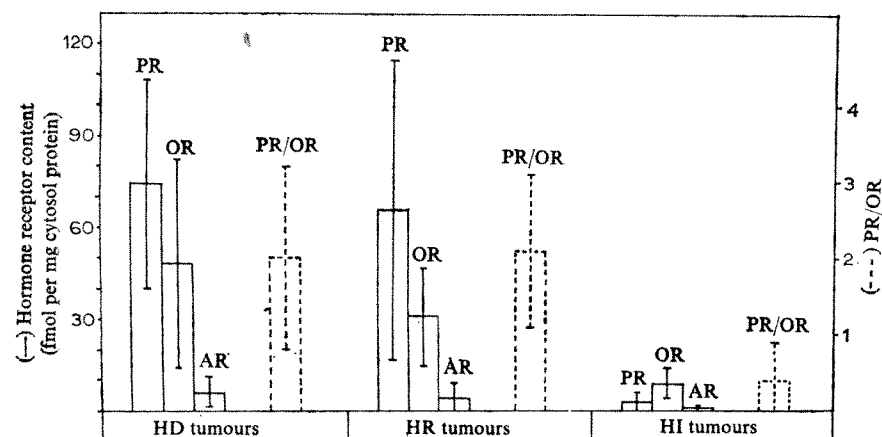

Fig. 4 Hormone-receptor contents of mammary tumours of GR mice. Receptors of progesterone (PR), oestrogen (OR) and androgen (AR). Receptor levels are expressed in fmol per mg cytosol protein. PR/OR, progesterone receptor/oestrogen receptor content ratio (mean ±s.d.). Hormone-dependent (HD), hormone-responsive (HR) and hormone-independent (HI) tumours.

contribute significantly to growth behaviour. Thus no essential difference in receptor levels was found between HD tumours and HR tumours (Fig. 4). Furthermore, tumours which became autonomous after serial transplantation in OPC mice usually still had significantly higher receptor contents than autonomous transplants derived from C mice (Fig. 1).

Variations in sampling also impose limitations on the significance of receptor assays. We often found marked differences in receptor values between different samples taken from one hormone-responsive tumour. If these reservations also apply to hormone-responsive breast cancers in human patients, they might contribute to the high percentage of clinically defined non-responding positives.

The finding that autonomous mammary tumours of GR mice generally had small but significant amounts of OR, but practically no PR, is of interest. Since in oestrogen target tissues the synthesis of PR depends on the action of oestrogen[16], it seems possible that in some autonomous mammary tumours of GR mice the lesion is at the nuclear level. For instance, this might be the case with some autonomous tumours of transplantation series I which had OR levels about equally high as some hor-

mone-responsive tumours of other lines (Fig. 1). This possibility has, however, to be investigated further.

1 Dao, T. L., A. Rev. Med , 23, 1–18 (1972).
2 Jensen, E. V., Block, G E , Smith, S , Kyser, K., and DeSombre, E. R., Natn. Cancer Inst. Monogr., 34, 55–70 (1971).
3 Maass, H., Engel, B , Hohmeister, H , Lehman, F , and Trams, G., Am. J. Obstet. Gynec., 113, 377–382 (1972).
4 Engelsman, E., Persijn, J. P , Korsten, C. B , and Cleton, F. J., Br. med. J., 2, 750–752 (1973).
5 McGuire, W. L., Carbone, P. P., Sears, M. E., and Escher, G. C., Estrogen Receptors in Human Breast Cancer (edit by McGuire, W L , Carbone, P. P., and Vollmer, E. P.), 1–7 (Raven, New York, 1975).
6 Boylan, E. S., and Wittliff, J L., Cancer Res., 35, 506–511 (1975).
7 Horwitz, K. B., McGuire, W. L , Pearson, O. H , and Segaloff, A., Science, 189, 726–727 (1975).
8 Persijn, J. P., Korsten, C B., and Engelsman, E., Br. med. J., 4, 503 (1975).
9 Sluyser, M., and Van Nie, R., Cancer Res , 34, 3253–3257 (1974).
10 Korsten, C B., and Persijn, J. P., Z. Klin. Chem. Biochem., 10, 502–508 (1972).
11 Korsten, C. B , Engelsman, E., and Persijn, J. P , Estrogen Receptors in Human Breast Cancer (edit. by McGuire, W L., Carbone, P. P., and Vollmer, E. P.), 93–105 (Raven, New York, 1975).
12 Philibert, D , and Raynaud, J P , Steroids, 22, 89–98 (1973).
13 Bonne, C., and Raynaud, J P , Steroids, 26, 227–232 (1975).
14 Lowry, O. H., Rosebrough, N J., Farr, A. L., and Randall, N. J , J. biol. Chem., 193, 265–273 (1951).
15 Kim, U., and Depowski, J., Cancer Res , 35, 2068–2077 (1975).
16 Freifeld, M L., Feil, P. D., and Bardin, C. W., Steroids, 23, 93–103 (1974).

Rotational diffusion of band 3 proteins in the human erythrocyte membrane

R. J. Cherry, A. Bürkli, M. Busslinger & G. Schneider

Eidgenössische Technische Hochschule, Laboratorium für Biochemie, ETH-Zentrum, CH8092 Zürich, Switzerland

G. R. Parish*

Eidgenössische Technische Hochschule, Institut für Zellbiologie, ETH-Hönggerberg, CH 8093 Zürich, Switzerland

Band 3 proteins rotate in the human erythrocyte membrane; rotation is almost certainly confined to an axis normal to the plane of the membrane and is characterised by a diffusion coefficient of the order of 1,000 s⁻¹. Spectrin does not restrict rotational motion of band 3 proteins.

MANY, and probably most, cell membranes contain lipids arranged in a bilayer configuration. According to currently popular concepts of membrane structure, 'intrinsic' or hydrophobic membrane proteins are essentially solutes in the fluid lipid bilayer[1]. As such they would be expected to undergo free diffusion—that is, diffusion the rate of which is determined only by the size of the protein and the 'viscosity' of the membrane. There is good evidence that this is the case for rhodopsin where Cone[2,3] has measured both lateral and rotational diffusion constants. In other instances there is qualitative evidence that proteins diffuse rather rapidly in the plane of the membrane (for reviews, see refs 4–6). On the other hand, at least one membrane protein, bacteriorhodopsin, is completely immobilised by protein–protein interactions[7–9]. At the present time, it is by no means clear to what extent proteins do diffuse freely and to what extent free diffusion is restricted by interactions with other components of the membrane, or conceivably of the cell—for example, microtubules or microfilaments.

To investigate this question, it is necessary to have available a technique which measures diffusion quantitatively and can be applied generally. In the case of rotational diffusion of membrane

proteins, rotational motion characterised by relaxation times at least as slow as microseconds has been both predicted and demonstrated[1,6]. This causes an experimental difficulty since such a time range is not adequately covered by established techniques of measurement. Indeed, investigation of the rotational diffusion of membrane proteins has only so far been possible with a very limited number of proteins which possess special spectroscopic properties[1,5–7].

To overcome this difficulty, we have developed a new technique for measuring slow rotational diffusion, which exploits the long lifetimes of triplet states of organic molecules[7,10,11]. Using the method of flash photolysis, triplet (or other long lived) states may be detected by transient absorbance changes after flash excitation. If the exciting light is plane polarised, the absorbance changes are in general dichroic. The rate of decay of dichroism gives information about the rate of rotational motion of the excited molecules.

Since most proteins do not naturally possess a suitable chromophore for such measurements, it is necessary to attach an artificial chromophore, or probe. Eosin (2,4,5,7-tetrabromofluorescein) is a favourable choice of probe since the four bromine atoms promote intersystem crossing. Covalent coupling to proteins may be achieved using the reactive isothiocyanate derivative (eosin–NCS). From studies with protein–eosin conjugates dissolved in viscous solutions, we have concluded that reliable measurements of rotational relaxation times in the microsecond–millisecond time range may be made using the flash photolysis method[10,11]. Here we report the first application of the method to proteins in a cell membrane.

The human erythrocyte membrane was chosen for these studies because the composition and properties of its proteins are relatively well characterised[12–14]. Major constituents are spectrin, a fibrous protein composed of two polypeptides of

*Present address: Department of Plant Sciences, University of Leeds, Leeds LS2 9JT, UK.

molecular weights 220,000 and 240,000, and band 3 (otherwise known as component *a*). Band 3 consists of hydrophobic proteins which span the membrane, have a molecular weight of about 90,000 and comprise about 24% of the total membrane protein[12]. They are implicated in anion and other transport functions[15-18]. The major sialoglycoprotein, glycophorin, also spans the membrane and has been extensively characterised[14]. It is thought that glycophorin in association with band 3 constitutes the membrane particles revealed by freeze-fracture electron microscopy[19,20].

Labelling of human erythrocytes with eosin–NCS

Red blood cells obtained from fresh or recently outdated human blood were washed three times with 310 mosM sodium phosphate buffer, *p*H 7.4. Packed cells (15 ml) were incubated with 3 mg eosin–NCS for 3 h at 22 °C. The cells were then washed twice more with isotonic buffer to remove any unreacted label and subsequently haemolysed in 40–50 volumes 20 mosM sodium phosphate buffer, *p*H 7.4. The ghosts were sedimented by centrifugation for 20 min at 20,000*g* and washed three to four times with 20 mosM buffer. All operations except the labelling step were carried out at 0–4 °C.

The amount of bound eosin was determined spectrophotometrically as described previously[10,11], after first solubilising ghosts with SDS. Protein was determined using the method of Lowry *et al.*[21]. Typically the ghosts contained 0.5–1.5 μg eosin per mg protein.

The proteins labelled by the above procedure were identified using SDS–polyacrylamide gel electrophoresis and by selective extraction experiments (details presented elsewhere). Briefly, SDS–polyacrylamide gel electrophoresis was carried out in a manner similar to that described by Fairbanks *et al.*[22]. Using gels of dimensions 0.85 × 8 cm loaded with 0.5 mg protein it was feasible to locate eosin fluorimetrically. Eosin fluorescence was located predominantly in the region of band 3. A small amount of eosin was also associated with bands 1 and 2 (spectrin). There were negligible amounts of eosin bound to other proteins. Using gels made from solutions containing 7.5% acrylamide (w/v), 2.73% *bis*-acrylamide (with respect to acrylamide), it was possible to separate glycophorin from band 3. In this way we could demonstrate that little or no eosin was bound to glycophorin.

These results were further confirmed using selective extraction procedures. Selective extraction of bands 1, 2 and 5 by incubation at low ionic strength in the presence of EDTA[22], demonstrated that less than 10% eosin was associated with these components. On the other hand, extraction with Triton X-100 (ref. 23), which has partial selectivity for band 3, released up to 80% eosin from the membrane. Finally the labelled ghosts were extracted with chloroform–methanol (2:1) as described by Hamaguchi and Cleve[24]. The amount of eosin extracted with lipids into the organic phase was negligible. A few per cent eosin appeared in the aqueous phase together with the membrane sialoglycoproteins.

The selective labelling which we have observed is in accord with other authors who find that only band 3 and sialoglycoproteins are available at the outer surface of the red cell[12, 25-27]. Since some of our label is associated with spectrin, we conclude that there is some permeation of eosin–NCS through the membrane. We found little labelling of glycophorin, and a similar observation was made by Cabantchik and Rothstein[15] with 4,4'-diisothiocyano-2,2'-stilbene disulphonic acid (DIDS). DIDS (added to intact cells) is found to bind almost exclusively to band 3 with less than 5% of the label associated with glycophorin.

The number of eosin molecules bound per protein is relevant to the possibility of energy transfer between eosins. For 1 μg eosin per mg total membrane protein, it may be estimated that the mole ratio eosin: band 3 would be 1:2 if all the eosin were bound to band 3 (assuming band 3 has a molecular weight of 90,000 and comprises 24% of the membrane protein).

Rotational diffusion measurements

Excitation of eosin at 540 nm by a laser light pulse of duration 1–2 μs and energy 100–200 mJ results in transient absorbance changes in the sample. Either positive absorbance changes (conveniently measured at 610–650 nm) due to triplet–triplet absorption or negative changes (measured at 500–525 nm) due to ground state depletion, may be detected. When the exciting source is plane-polarised, the absorbance changes are dichroic provided rotational motion is slower than the instrumental response (2–3 μs)[10,11].

Transient absorption changes from eosin-labelled ghosts are shown in Fig. 1. The better signal-to-noise ratio of the ground state depletion signal is due to the higher extinction coefficient of the singlet–singlet transition compared with the triplet–triplet transition. Both absorption and depletion signals demonstrate the existence of a rather long lived dichroism. All quantitative evaluation of the data was carried out with depletion signals.

The data were analysed by calculating the anisotropy factor *r*(*t*) given by

$$r(t) = (A_{pa}(t) - A_{pe}(t))/(A_{pa}(t) + 2A_{pe}(t)) \qquad (1)$$

where $A_{pa}(t)$ and $A_{pe}(t)$ are respectively the absorbance changes at time *t* for light polarised parallel and perpendicular to the exciting flash. The time dependence of *r* is shown on a semi-

Fig. 1 Transient absorbance changes from eosin-labelled ghosts. Membranes suspended in 20 mosM phosphate buffer, *p*H 7.4, temperature 22 °C. Eosin concentration 7 μM, 0.9 μg eosin per mg protein. Before measurement, samples were bubbled with argon for 10 min to displace oxygen. *a*, Triplet–triplet absorption measured at 650 nm; *b*, ground state depletion measured at 525 nm. In each case the stronger signal is obtained with the measuring beam polarised parallel to the exciting flash, the weaker signal for perpendicular polarisation. The flash photolysis apparatus was described in ref. 11.
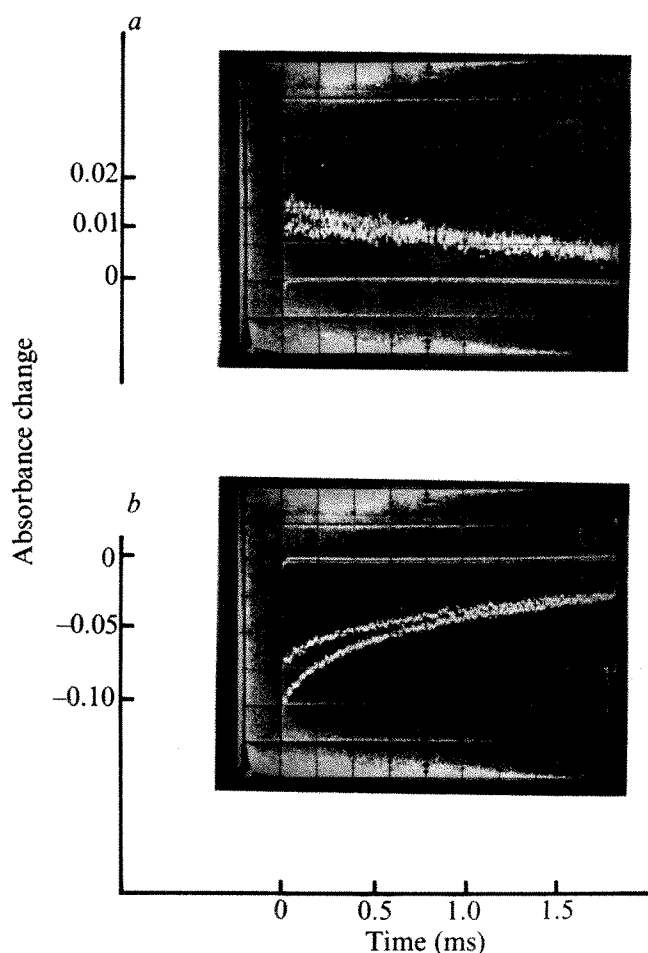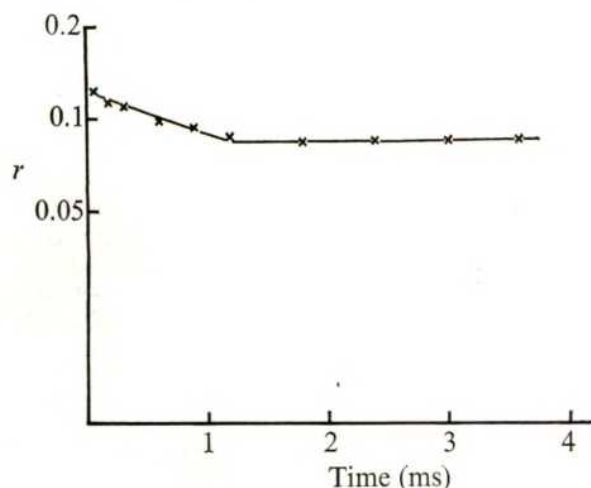
Fig. 2 Time dependence of the anisotropy parameter r calculated from depletion signals similar to those shown in Fig. 1b. For more accurate calculation the difference between the parallel and perpendicularly polarised components is displayed directly using the differential input of the oscilloscope amplifier.

logarithmic plot in Fig. 2. It is clear that the decay of dichroism contains at least two components. Within the experimental uncertainty $r(t)$ is independent of time at long times so that the observed time dependence of r takes the form

$$r(t) = r_1 \exp(-t/\varphi_1) + r_2 \qquad (2)$$

Replotting $\log(r(t) - r_2)$ against t gives $\varphi_1 \approx 0.5$ ms. (The data are not sufficiently accurate to determine φ_1 precisely; conservatively, we consider the above value to be correct within a factor of 2.)

Interpretation of data

Consideration of the magnitude of the flash photolysis signals leads us to conclude that a high proportion of eosin molecules in the sample must contribute to the observed dichroism. Since we have shown that eosin is mainly bound to band 3, we can relate the dichroism data to the rotational motion of these proteins.

Equations which evaluate $r(t)$ for an irregular body in an anisotropic medium are not currently available. Various authors have, however, solved the problem of the time dependence of r for an irregular body in isotropic solution[28-31]. If we assume that the rotational motion of band 3 proteins may be characterised by two diffusion constants, D_{pa} for rotation about an axis normal to the membrane and D_{pe} for rotation about either of the two perpendicular axes, then the problem formally becomes equivalent to the rotation of a body with an axis of symmetry immersed in an isotropic medium. We can therefore use existing theoretical treatments to evaluate our data. An additional simplification arises with depletion signals because the same transition moment is used for excitation and measurement. (The lowest absorption band of eosin is a single transition and triplet–triplet absorption is relatively small at 525 nm (refs 10 and 11).) The theoretical expression for $r(t)$ then becomes

$$r(t) = \sum_{i=1}^{3} A_i \exp(-E_i t) \qquad (3)$$

where $A_1 = 6/5 \sin^2\theta \cos^2\theta$, $A_2 = 3/10 \sin^4\theta$, $A_3 = 1/10(3\cos^2\theta - 1)^2$, $E_1 = 5D_{pe} + D_{pa}$, $E_2 = 2D_{pe} + 4D_{pa}$, $E_3 = 6D_{pe}$, and θ is the angle between the transition moment and the axis of symmetry. (The theoretical expressions given in refs 28–31 all become equivalent for the case considered here.)

We consider that there are two main interpretations of the two experimentally observed components in the decay of $r(t)$. Since band 3 proteins seem to offer different labelling sites on either side of the membrane[12,25-27], it is likely that they have a negligibly small rate of rotation about an axis in the plane of the membrane. The existence of a carbohydrate moiety on these proteins, which is only detectable on the outside surface, points to the same conclusion[16,32,33]. Thus we anticipate that $D_{pe} \approx 0$ and equation (3) then gives

$$r(t) = A_1 \exp(-D_{pa}t) + A_2 \exp(-4D_{pa}t) + A_3 \qquad (4)$$

Thus, the time-independent component of the experimental $r(t)$ curve may be identified with A_3 and is due to the negligibly small value of D_{pe}. ($r(t)$ will of course eventually go to zero due to rotation of the whole erythrocyte ghost; the relaxation time of this process is, however, of the order of seconds.) The measured time constant φ_1 is then related to D_{pa}; the experimental accuracy is, however, clearly insufficient to detect the two expected components.

An alternative explanation of the data is that the eosin-binding sites are heterogeneous. (Note, however, that the amount of eosin associated with spectrin is too small to account for one or other component.) In particular, it cannot be ruled out that there is some aggregation of band 3 proteins in the membrane. In this case r_2 could correspond to aggregates which are too large to have an appreciable rotation on the time scale of the experiment.

If the first explanation is correct, we could in principle determine θ (which is likely to be a mean value), since $r_2 = A_3$. This is, however, dangerous since even if protein aggregation is not the main reason for the time-independent dichroism, it may well make a contribution. (The observation that the relative magnitudes of r_1 and r_2 vary from one preparation to another makes us suspicious that this may be the case.) We therefore prefer not to attempt to determine θ (and therefore the coefficients A_1 and A_2) and to regard φ_1 as an unknown combination
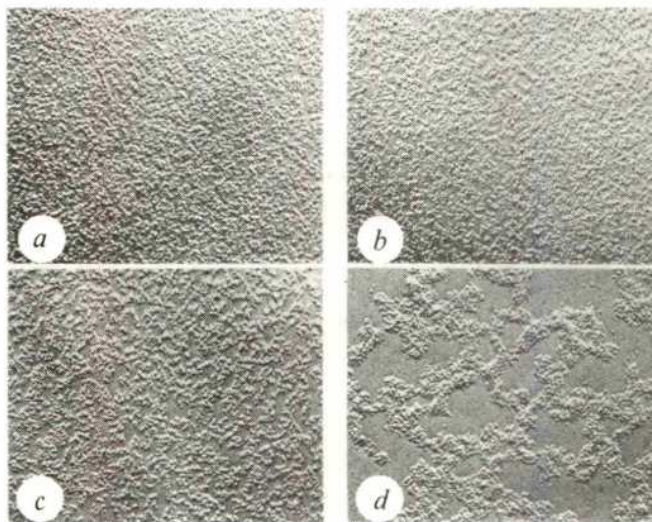
Fig. 3 Inner fracture faces (IFF) of ghosts before and after spectrin depletion ($\times 36,000$). a, Unlabelled, untreated, pH 7.6; b, eosin-labelled, untreated, pH 7.6; c, eosin-labelled, spectrin-depleted, pH 7.6; d, eosin-labelled, spectrin-depleted, pH 5.4. Eosin labelling as in text. Spectrin was depleted by incubating 1 part packed ghosts with 10 parts 20 mM phosphate buffer, pH 8.5, (containing 0.5 mM NaN$_3$) for 18 h at 37 °C (ref. 43). Samples were maintained at 0–4 °C at all times except during labelling and spectrin removal. For electron microscopy, samples were mounted on precooled gold disks and frozen in liquid Freon. Freeze fracturing was carried out in a Balzer's apparatus at a temperature of −100 or −105 °C. Specimens (without etching) were replicated with Pt–C and backed with C, both evaporated from electron beam guns. Replicas were examined in a Phillips 200 electron microscope.

of $1/D_{pa}$ and $1/4D_{pa}$. On this basis a value of 0.5 ms for φ_1 implies D_{pa} is in the range 500–2,000 s^{-1}.

Band 3 is probably involved in anion transport across the erythrocyte membrane[15,16,34]. Lepke *et al.*[35] find that sulphate exchange is completely inhibited when 1.7×10^6 molecules of the irreversible inhibitor dihydro 4,4′-diisothiocyano-1,2-diphenyl-ethane-2,2′-disulphonic acid (H$_2$DIDS) are bound to band 3. This figure implies a turnover number for chloride exchange of about 2×10^4 per s per site if all these sites are involved in transport. If only a portion of these sites are responsible for transport then the turnover number must be higher. Thus a rotation time in the order of tens of microseconds or faster would be required if transport occurred by a rotational mechanism. Irrespective of any particular interpretation, the present experiments do not detect any component of rotational motion of band 3 in this order of magnitude. Thus if band 3 is responsible for anion transport, a mechanism involving rotation of the whole protein is ruled out. This is hardly a surprising conclusion; as discussed above, arguments based on the structure of band 3 also indicate that rotation across the membrane is improbable[16,25]. Nevertheless, the present studies demonstrate how rotation measurements may be useful in elucidating transport mechanisms in other systems where the outcome is not so easily anticipated.

'Viscosity' of the membrane

Whatever the precise interpretation of the data, a component of rotational motion of band 3 proteins of the order of 0.5 ms is clearly present. The only existing quantitative determination of the rotation of a membrane protein is Cone's measurement with rhodopsin[2], which yielded a relaxation time of 20 μs at 20 °C. Rotation of band 3 proteins in the human erythrocyte membrane clearly is markedly slower.

We found that addition of glycerol to a concentration of 60% to the membrane suspension had no detectable effect on φ_1. Since this addition increased the medium viscosity by a factor of 10, we may conclude that the rotation of band 3 proteins is a function of the membrane viscosity. If we regard band 3 proteins as cylinders of radius a, then the friction coefficient for rotation about the axis of symmetry is given by (ref. 36)

$$f = 4\pi a^2 h \eta \tag{5}$$

where h is the depth of the cylinder immersed in the membrane and η is the membrane viscosity.

Fig. 4 Effect of spectrin removal on the time dependence of the anisotropy parameter r. Temperature 22 °C. ×, Control sample, pH 7.6; ●, spectrin-depleted ghosts (prepared as described by Fairbanks *et al.*[22]), pH 7.6; △, spectrin-depleted ghosts, pH 5.4. The lower initial value of r for the control sample is due to experimental factors and has no special significance. Essentially similar results were obtained when spectrin was depleted using the procedure of Elgsaeter and Branton[43].
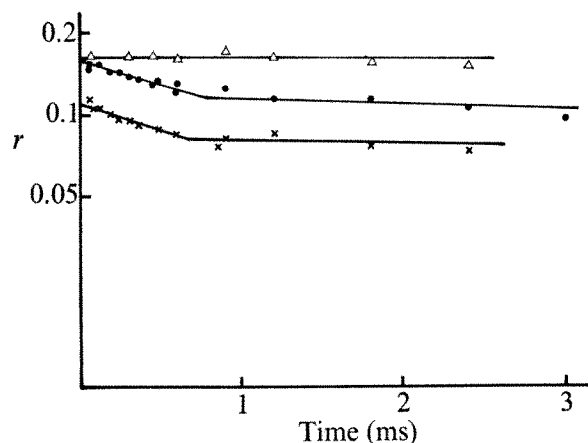
From the Einstein relation[37]

$$D = kt/f \tag{6}$$

we obtain

$$D_{pa} = kt/4\pi a^2 h \eta \tag{7}$$

If we take a as the radius of the particles revealed by freeze-fracture electron microscopy (40 Å) and h as the thickness of the hydrophobic core of the membrane (40 Å), then a value of $\varphi_1 = 0.5$ ms implies $\eta = 25 - 100P$ (remembering that the respective contributions of $1/D_{pa}$ and $1/4D_{pa}$ to φ_1 have not been determined). This is considerably higher than previous estimates of the viscosity of the erythrocyte membrane based on measurements of methanol diffusion[38] and fluorescence polarisation[39,40]. It may indicate that the 'microviscosity' experienced by small molecules dissolved in the hydrocarbon region of the lipid bilayer is quite different from that experienced by a protein which may interact with the whole length of the lipid molecules.

Probably the replication procedure used in freeze-fracture electron microscopy leads to an overestimate in the size of the particles[41,42]. If the true radius of band 3 proteins is less than 40 Å, then the calculated viscosity will be increased to still higher values.

Interaction between band 3 and spectrin

The above discussion of membrane viscosity only has validity if band 3 proteins undergo free diffusion in the membrane. Apart from the possibility of self-association of these proteins discussed above, possible interactions with other components of the membrane have also to be considered. In particular, there is evidence that spectrin restricts lateral diffusion of intramembrane particles[43]. It has been proposed that these particles contain both glycophorin and band 3 and are physically linked to spectrin[19].

We have investigated the effects of spectrin removal on rotational diffusion of band 3 proteins. We have depleted eosin-labelled ghosts of spectrin using the procedure of incubation at low ionic strength described by Elgsaeter and Branton[43]. Figure 3 shows freeze-fracture electron micrographs of unlabelled ghosts, eosin-labelled ghosts and spectrin-depleted ghosts, all at pH 7.6, together with spectrin-depleted ghosts at pH 5.4. The distribution of particles is little different in the labelled ghosts compared with the control, demonstrating that the labelling procedure in itself does not cause any marked aggregation. In agreement with Elgsaeter and Branton[43], the distribution of particles in spectrin-depleted ghosts is little altered at pH 7.6 but massive aggregation occurs on lowering the pH to 5.4.

Spectrin is most efficiently and selectively removed by incubation at low ionic strength in the presence of EDTA. Using the procedure of Fairbanks *et al.*[22] we have also prepared spectrin-depleted ghosts in this manner. SDS–polyacrylamide gels confirmed that most of bands 1, 2 and 5 were selectively released from the membrane. When examined by electron microscopy it was clear that considerable vesiculation of the membranes had occurred, as noted by other authors[44]. The effects of pH on particle aggregation was, at far as could be ascertained, essentially similar to that shown in Fig. 3.

Spectrin-depleted ghosts prepared by both methods outlined above were examined by flash photolysis. Essentially similar results were obtained in both cases; one set of data is shown in Fig. 4.

Inspection of Fig. 4 clearly shows that there is no detectable rotational motion in spectrin-depleted ghosts at pH 5.4. This is entirely reasonable in view of the strong aggregation of particles in these conditions as seen in Fig. 3. Large aggregates would be expected to have negligible rotation on the time scale of the experiment. This correlation between rotational motion and the state of aggregation of the particles further demonstrates that we are making valid measurements of rotational motion of band 3 proteins (or alternatively, provides further evidence that the particles contain band 3 proteins).

The most interesting conclusion which arises from the results in Fig. 4 is that there is no detectable difference in rotational motion between spectrin-depleted ghosts and control ghosts when the *p*H is maintained at 7.6. In other words, there is no evidence that spectrin restricts the rotational diffusion of band 3 proteins. Although this finding does not completely rule out a direct linkage between band 3 and spectrin, it does place severe restrictions on the type of attachment which can be envisaged. Only a linkage which permits free rotation of band 3 about an axis normal to the plane of the membrane is compatible with the present measurements.

On the other hand, we consider that a model in which there is no direct interaction between band 3 and spectrin accounts equally well for all the mobility data. Suppose that band 3 proteins protrude from the membrane in the spaces between the spectrin fibres. It follows that lateral diffusion over large distances would be prevented by the spectrin network, whereas rotational diffusion would be unaffected. The enhanced aggregation of membrane particles in spectrin-depleted ghosts at low *p*H (ref. 43) would be explained by such a model. Further, proteins entrapped within the spectrin network would probably follow any redistribution of the spectrin molecules themselves. The clustering of anion sites[45] and concanavalin A receptors[19] on the outer membrane surface which is observed when spectrin is aggregated may be explained in this way.

Peters *et al.*[46] labelled proteins (principally band 3) with fluorescein, bleached one half of a single ghost and looked for the spread of fluorescence from the unbleached to the bleached half. No spreading was detected, placing an upper limit on the lateral diffusion constant of 3×10^{-12} cm^2 s^{-1}. For free diffusion such a low value would imply a rotational relaxation time of the order of 50 ms, a value clearly incompatible with our 0.5 ms component of rotational motion. This again implies that lateral diffusion is restricted and is entirely compatible with a model of band 3 proteins trapped in the spectrin net.

Our results emphasise the importance of clearly distinguishing local diffusion from diffusion over large distances. The fact that the latter is not observed does not necessarily imply that proteins are immobilised. If the membrane is 'compartmentalised' then rapid diffusion over small distances may still occur and in turn may have functional significance.

We thank Professor G. Semenza for encouragement, Dr K. Razi Naqvi for discussions and the Swiss NSF for support.

1 Singer, S. J., and Nicolson, G. L., *Science*, 175, 720–731 (1972).
2 Cone, R. A., *Nature new Biol.*, 236, 39–43 (1972).
3 Poo, M., and Cone, R. A., *Nature*, 247, 438–441 (1974).
4 Edidin, M., *A. Rev. Biophys. bioengng*, 3, 179–301 (1974).
5 Cherry, R. J., *FEBS Lett.*, 55, 1–7 (1975).
6 Cherry, R. J., *Biological Membranes*, 3 (edit. by Chapman, D. and Wallach, D. F. H.), 47–102 (Academic, London and New York, 1976).
7 Razi Naqvi, K., Gozalez-Rodriguez, J., Cherry, R. J., and Chapman, D., *Nature new Biol.*, 245, 249–251 (1973).
8 Blaurock, A. E., and Stoeckenius, W., *Nature new Biol.*, 233, 152–154 (1971).
9 Henderson, R., and Unwin, P. N. T., *Nature*, 257, 28–32 (1975).
10 Cherry, R. J., Cogoli, A., Oppliger, M., Schneider, G., and Semenza, G., *Biochemistry* (in the press).
11 Cherry, R. J., and Schneider, G., *Biochemistry* (in the press).
12 Steck, T. L., *J. Cell Biol.*, 62, 1–19 (1974).
13 Juliano, R. L., *Biochim. biophys. Acta*, 300, 341–378 (1973).
14 Tomita, M., and Marchesi, V. T., *Proc. natn. Acad. Sci. U.S.A.*, 72, 2964–2968 (1975).
15 Cabantchik, Z. I., and Rothstein, A., *J. membr. Biol.*, 15, 207–226 (1974).
16 Ho, M. K., and Guidotti, G., *J. biol. Chem.*, 250, 675–683 (1975).
17 Lin, S., and Spudich, J. A., *J. biol. Chem.*, 238, 5778–5783 (1974).
18 Brown, P. A., Feinstein, M. B., and Sha'afi, R. I., *Nature*, 254, 523–525 (1975).
19 Pinto da Silva, P., and Nicolson, G. L., *Biochim. biophys. Acta*, 363, 311–319 (1974).
20 Guidotti, G., *A. Rev. Biochem.*, 41, 731–752 (1972).
21 Lowry, O. H., Rosenbrough, N. J., Farr, A. L., and Randall, R. J., *J. biol. Chem.*, 193, 265–275(1951).
22 Fairbanks, G., Steck, T. L., and Wallach, D. F. H., *Biochemistry*, 10, 2606–2616 (1971).
23 Yu, J., Fischman, D. A., and Steck, T. L., *J. supramolec. Struct.*, 1, 233–248 (1973).
24 Hamaguchi, H., and Cleve, H., *Biochem. biophys. Res. Commun.*, 47, 459–464 (1972).
25 Bretscher, M. S., *Science*, 181, 622–629 (1973).
26 Morrison, M., Mueller, T. J., and Huber, C. T., *J. biol. Chem.*, 249, 2658–2660 (1974).
27 Shin, B. C., and Carraway, K. L., *Biochim. biophys. Acta*, 345, 141–153 (1974).
28 Tao, T., *Biopolymers*, 8, 609–632 (1969).
29 Belford, G. G., Belford, R. L., and Weber, G., *Proc. natn. Acad. Sci. U.S.A.*, 69, 1392–1393 (1972).
30 Ehrenberg, M., and Rigler, R., *Chem. phys. Lett.*, 14, 539–544 (1972).
31 Chuang, T. J., and Eisenthal, K. B., *J. chem. Phys.*, 57, 5094–5097 (1972).
32 Tanner, M. J. A., and Boxer, D. H., *Biochem. J.*, 129, 333–347 (1972).
33 Steck, T. L., and Dawson, G., *J. biol. Chem.*, 249, 2135–2142 (1974).
34 Passow, H., Fasold, H., Zaki, L., Schuhmann, B., and Lepke, S., *Biomembranes: Structure and Function* (edit. by Gardos, G., and Szász, I.), 197–214 (Publishing House of the Hungarian Academy of Sciences, Budapest, 1975).
35 Lepke, L., Fasold, H., Pring, M., and Passow, H., *J. membr. Biol.* (in the press).
36 Saffman, G., and Delbrück, M., *Proc. natn. Acad. Sci. U.S.A.*, 72, 3111–3113 (1975).
37 Einstein, A., *Ann. Physik*, 17, 549–562 (1905).
38 Solomon, A. K., *Biochim. biophys. Acta*, 373, 145–149 (1974).
39 Rudy, B., and Gitler, C., *Biochim. biophys. Acta*, 288, 231–236 (1972).
40 Aloni, B., Shinitzky, M., and Livne, A., *Biochim. biophys. Acta*, 348, 438–441 (1974).
41 Weinstein, R. S., *Red Cell Membrane Structure and Function* (edit. by Jamieson, G.A.), 36–76 (J. B. Lippincott, Philadelphia, 1969).
42 Nermut, M. V., *Proc. R. microscop. Soc.*, 10, 104 (1975).
43 Elgsaeter, A., and Branton, D., *J. Cell Biol.*, 63, 1018–1030 (1974).
44 Marchesi, V. T., and Steers, E., *Science*, 159, 203–204 (1968).
45 Nicolson, G. L., and Painter, R. G., *J. Cell Biol.*, 59, 395–406 (1973).
46 Peters, R., Peters, J., Tews, K. H., and Bähr, W., *Biochim. biophys. Acta*, 367 282–294 (1974).

letters to nature

Optical behaviour of HDE226868 during a Cyg X-1 X-ray transition

As part of regular monitoring of variable X-ray sources with the University College London–Mullard Space Science Laboratory X-ray instrumentation on the Copernicus satellite, we have observed Cyg X-1 continuously for 12 d between 22 October and 3 November, 1975. The X-ray flux was noted to be more variable than it had been during previous monitoring runs, when the source had been in its low state[1], and, in particular, a marked increase in flux occurred 8 d into the observing run (~ October 29.5 UT, 1975). After this date, the mean 3–8-keV flux increased steadily, and by 3 November had risen by 30% compared with its level during the first part of the observation. The X-ray flux data, integrated in time blocks of 30 min, are shown in Fig. 1b. In Fig. 1a we show the spectral hardness ratio (defined as the 5.5–8.0-keV flux divided by the 3.0–5.5-keV flux) integrated over intervals of 0.5 d. This quantity decreased

progressively after 30 October, until by 3 November it was down by an amount equivalent to a drop of 0.1 in the power law spectral index of the source.

Holt *et al.*[2] and Grindlay *et al.*[3] have also noted an increase in the X-ray flux of Cyg X-1 at about this time. These two groups found the sources to be up to five times stronger than the

Table 1 Best-fit parameters and χ^2

Interval	χ^2 per degree of freedom	Mean ($m_{\rm B}$)	Semi-amplitude ($m_{\rm B}$)
A	23.9	9.644	0.025
B	18.6	9.660	0.018
C	28.4	9.661	0.022
D	32.5	9.650	0.018
E	18.8	9.649	0.016
F	7.7	9.640	0.028
G	19.1	9.650	0.020
All 1975 data	22.0	9.652	0.020
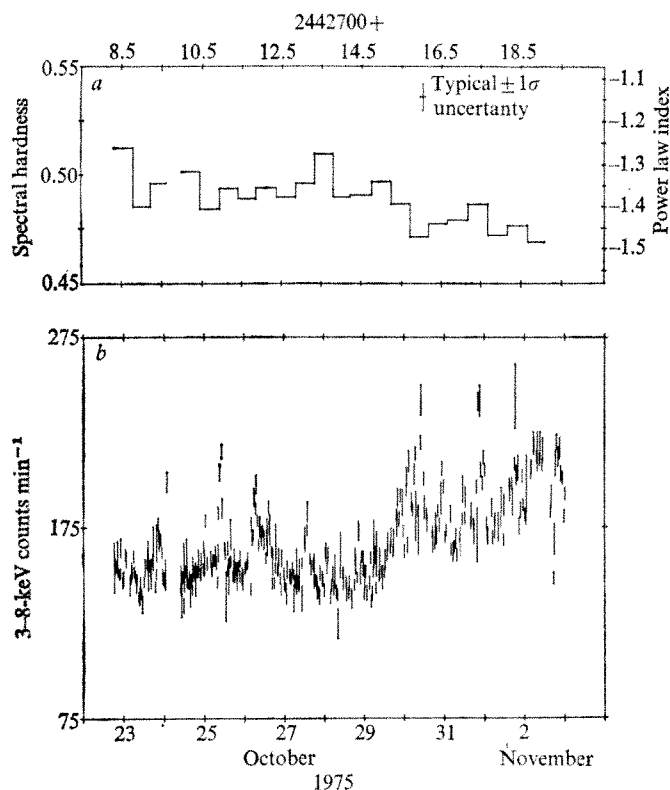
Fig. 1 Copernicus X-ray data for Cyg X-1 during October–November 1975. *b*, X-ray flux binned into 30-min intervals; *a*, spectral hardness ratio in 0.5-d bins. The error bars correspond to $\pm 1\sigma$ in each plot and there is a systematic uncertainty of up to 0.1 in the power law index indicated on the right-hand ordinate of *b*.

level seen earlier in October, a maximum equivalent to $\simeq 1,500$ Uhuru counts being seen on 6 November[3]. Continued observation with the Ariel V satellite[4,5] has shown that Cyg X-1 maintained a relatively high flux level until late February, when the intensity dropped by a factor of three[5].

The blue-band optical magnitude of HDE226868, the spectroscopic binary BO1b star which is the optical counterpart of Cyg X-1, was monitored between July and December 1975 from the Spanish Sierra Nevada. The data are shown in Fig. 2, and are compared with the X-ray behaviour of Cyg X-1 as measured by Copernicus and inferred from the Ariel V and ANS data reported in refs 2 and 3. The crosses represent nightly mean blue magnitudes and have a typical standard error of $\pm 0.003 m_B$.

We note, first, that the mean B-band flux reaches a local maximum close to the time of the X-ray transition. This may be coincidental, however, since such maxima are not unprecedented and have not previously been correlated with any known X-ray event. Of more significance is the detailed form of the blue light curve about the time of the X-ray transition. As observed in previous years, the average blue light curve

for the 1975 data is typified by a distorted double sine wave per orbital period[6–8]. To quantify the light curve distortion, we have fitted a double sine wave to the individual three-orbit intervals marked in Fig. 2, allowing the amplitude and mean magnitude to vary so as to minimise χ^2. The phase of the double sine wave was taken to be the mean phase obtained by fitting all of the 1975 data.

The best-fit parameters and the corresponding reduced χ^2 for each interval are shown in Table 1 and light curves for the intervals about the time of the X-ray transition are shown in Fig. 3. The data points and the best-fit double sine wave for interval E are shown in Fig. 3*a*. This light curve is typical of the data from July to October 1975, and the large values of χ^2 for intervals A to E indicate the relatively large departures from a double-sine-wave light curve.

A change in the character of the light curve was observed, starting on the night of 25/26 October, as shown in Fig. 3*b*. On this night the B-band flux was abnormally faint, and then increased to an abnormally high value two nights later. No light variations were then seen over the following three nights.
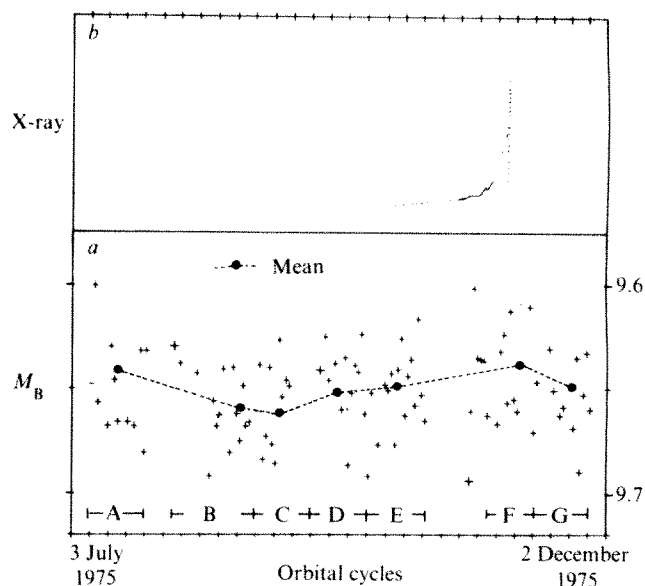
Fig. 2 *a*, Blue-band magnitude of HDE226868 from July 1975. The filled circles represent means (uncertainty $\sim 0.004 m_B$) over the intervals marked (~ 3 orbits). The X-ray flux from the present work (——) and from refs 3 and 4 (— —) are shown in *b*. Times of superior conjunction of the X-ray source are indicated on the abscissae.

During the interval immediately after the X-ray turn-on, regular light variations were again observed, as indicated in Fig. 3*c*, which shows data and the best-fit double sine wave for interval F. The degree of departure from the best-fit curve was, however, considerably reduced, as demonstrated by the low value of χ^2 for this interval. (Note that the six data points shown in Fig. 3*b* are not included in interval F.) Distortion of
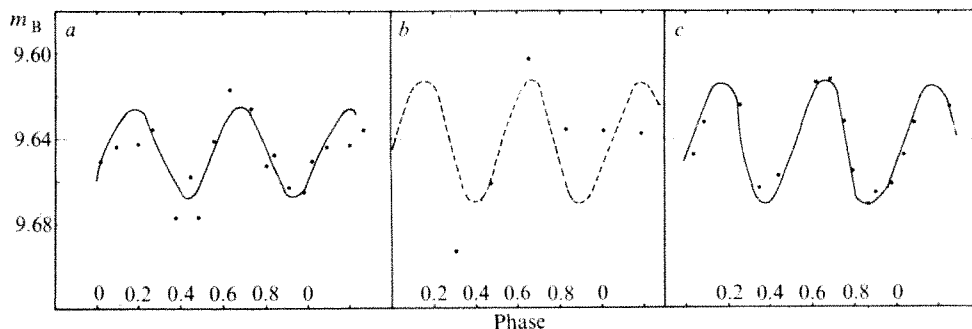
Fig. 3 *a* and *c*, Blue-band data folded with the 5.6-d orbital period. *b*, Data (unfolded) for the 6 d spanning the transition in X-ray flux noted in the text. Formal errors of measurement are $\pm 0.003 m_B$. The best-fit double sine curves are indicated for the data in *a* and *c*. The best-fit curve to the data shown in *c* is superimposed on the unfolded data in *b*, for reference. *a* (Interval E), JD 2442684–698; *b*, JD 2442711–716; *c* (Interval F), JD 2442717–736.

the light curve became more pronounced again towards the end of the 1975 data, as indicated by the higher value of χ^2 for interval G. It is important to note that the behaviour described above is unprecedented in a total of 400 nights of optical monitoring of the source distributed through the past 4 yr.

As the unique change in the optical behaviour of HDE226868 coincides to within one orbital period with the onset of the transition of Cyg X-1 to its high state, it is highly probable that the X-ray and optical events are related. Our immediate conclusion is that the optical variability is influenced by the condition of the X-ray source, and if ellipsoidal-like variations, as discussed by Hutchings[9] and Avni and Bahcall[10], do occur, then at times they are insignificant by comparison with this other mechanism.

E. N. WALKER

Royal Greenwich Observatory,
Herstmonceux Castle,
Hailsham, East Sussex, UK

G. D. BROWNLIE
K. O. MASON
P. W. SANFORD

Mullard Space Science Laboratory,
University College London,
Holmbury St Mary,
Dorking, Surrey, UK

A. R. QUINTANILLA

Instituto de Astrofisica de Andalucia,
University of Granada, Spain

1 Sanford, P. W., Ives, J. C., Bell-Burnell, S. J., Mason, K. O., and Murdin, P., *Nature,* **256**, 109 (1975).
2 Holt, S. S., Kaluzienski, L. J., Boldt, E. A., and Serlemitsos, P. J., *IAU Circ.,* No. 2863 (1975).
3 Grindlay, J., Schreier, E., den Boggende, A., and Brinkmann, A., *IAU Circ.,* No. 2863 (1975).
4 Holt, S. S., Kaluzienski, L. J., Boldt, E. A., and Serlemitsos, P. J., *IAU Circ.,* No. 2879 (1975).
5 Holt, S. S., Kaluzienski, L. J., Boldt, E. A., and Serlemitsos, P. J., *Nature,* **261**, 213 (1976).
6 Walker, E. N., *Mon. Not. R. astr. Soc.,* **160**, 9P (1972).
7 Walker, E. N., and Rolland, Q. A., *Mon. Not. R. astr. Soc.,* **169**, 247 (1974).
8 Walker, E. N., Rolland, Q. A., and Buck, C. S., *Mon. Not. R. astr. Soc.* (in the press).
9 Hutchings, J. B., *Astrophys. J.,* **188**, 341 (1974).
10 Avni, Y., and Bahcall, J. N., *Astrophys. J.,* **197**, 657 (1975).

Suspected globular clusters in the Fornax I cluster of galaxies

HERE we report the existence of a significant clustering of faint objects around three galaxies in the Fornax I cluster of galaxies. The nature of these objects is uncertain, although on the scant evidence available, we believe at least some of them to be unresolved globular clusters. We base our claim on the examination of a single, direct photographic plate taken by Dr R. D. Cannon at the prime focus of the Anglo–Australian telescope on the night of January 18, 1975. The plate, AAT609, is an unsensitised IIaO, exposed for 20 min through a GG 385 filter, and covers an area of $1° \times 1°$ centred on RA 3 h 36 min, dec. $-35.5°$ (1950.0). The smallest images on the plate are slightly elliptical, owing to slight trailing in RA, and have dimensions $1.5'' \times 2.0''$. In the course of assigning morphological types to the brightest galaxies in the field, we found a noticeable increase in the number of faint objects in the vicinity of the galaxies NGC1374, NGC1379 and NGC1399. These three galaxies are all elliptical and are among the nine brightest on the plate. Of the remaining six bright galaxies, two (NGC1389 and NGC1404) are elliptical and four (NGC1375, NGC1380, NGC1381 and NGC1387) are lenticular; none of these six galaxies exhibits any significant enhancement in the number of adjacent field objects.

In Fig. 1, we show the marginal distributions in RA and dec. of the faint objects around the three galaxies, and in
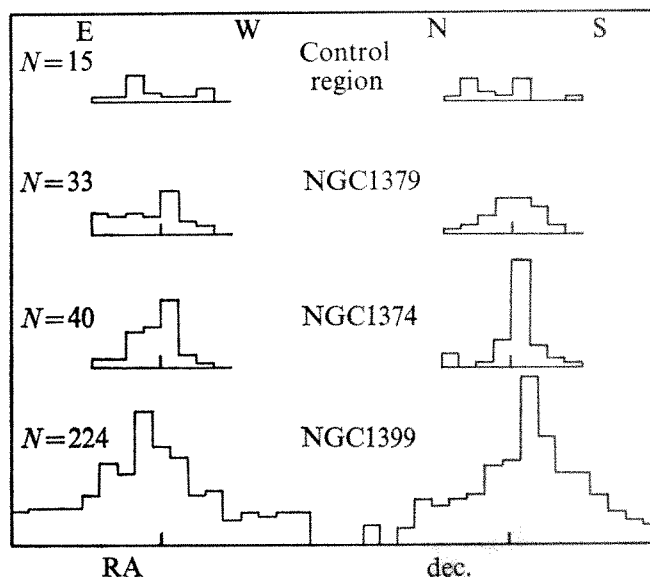
Fig. 1 Marginal distributions of faint objects around three galaxies in Fornax I, and in a control region. The ordinates are in the same arbitary units for all the histograms. N represents the total number of faint objects in each area measured. The widths of the histograms for NGC1399 are both 10.7', those for the remaining regions are 5.0'.

a control area centred on a star at RA 3 h 34 min, dec. $-35.3°$, well removed from any bright galaxy. The counts were made visually with the aid of a low-powered microscope and a rectangular grid (of unit size $18.9'' \times 18.9''$ on the plate), and the positions of all objects were recorded. The area measured round NGC1399, the brightest of the three galaxies, was $10.7' \times 10.7'$, whereas the corresponding areas for the other two galaxies and the control region were $5.0' \times 5.0'$. Those objects judged to be faint were selected solely on the criterion of image size and not of image structure. They were estimated to be within 1 mag of the limiting magnitude of the plate. In all, 224 were found around NGC1399, 40 around NGC1374 and 33 around NGC1379. For comparison, the number of faint objects in the control region was 15.

The density distribution of faint objects around NGC1399, derived from ring counts (Fig. 2) has been fitted with a truncated isothermal distribution assuming spherical symmetry. The asymptotic value of the density of $3.4 \times 10^3 \pm 1.4 \times 10^3$ (s.e.) per degree^2 is slightly higher than the density of 2.2×10^3 per degree^2 found in the control region. In view of the differences in the dispersion of the marginal distributions in RA and dec. (Fig. 1), we thought it necessary to check the assumption of spherical symmetry. We have fitted the observations with a bivariate Gaussian distribution and found major and minor axes of $154''$ and $133''$ (that is, an ellipticity of 0.13), and that the position angle of the major axis is $105°$. This weakly confirms the impression given to the eye that the cluster of faint objects around NGC1399 is slightly elongated in the direction of the major axis of the galaxy (position angle $115°$). The centroid of the distribution however, lies $28''$ to the SE of the centre of NGC1399, and there is a relative paucity of objects in the NE sector of the galaxy. The NGC1374 system seems symmetrical and well defined. The smaller number of objects around NGC1379 inhibits any assessment of the symmetry of the distribution.

The average density of all stellar images on the plate is 4×10^3 deg^{-2}. If we tentatively identify these with foreground stars within our Galaxy, then we may extrapolate an approximate value for the limiting magnitude of the plate from tables of star numbers given by Allen[1] (with the appropriate correction for magnitude scale errors given by
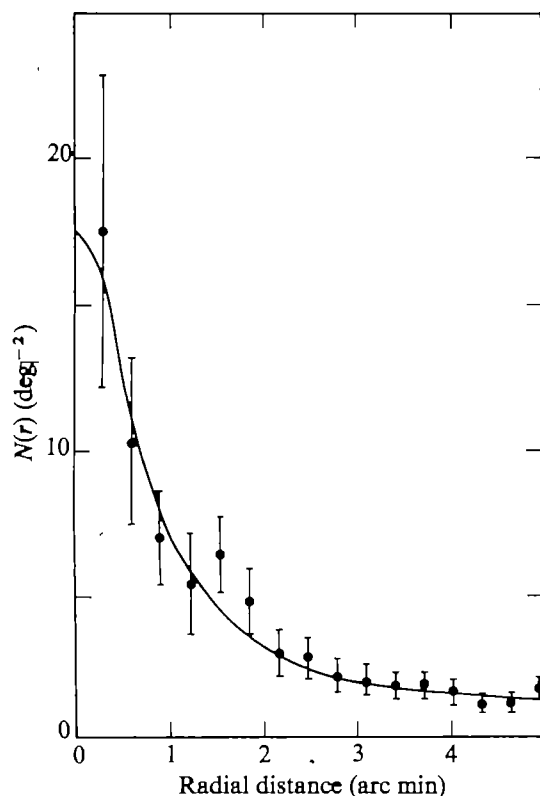
Fig. 2 The radial density distribution (number arc min^{-2}) of the faint objects around NGC1399 (dots, with corresponding error bars). The best isothermal fit is shown by the continuous curve.

Arp[7]). For the galactic latitude of Fornax I, $b^{\text{II}} \sim -54°$, the above density corresponds to a limiting B magnitude of 22.5±1.5 mag (estimated error). Although plate AAT609 was not calibrated, another estimate of its limiting magnitude may be obtained from a comparison with plate AAT610 taken on the same night. This is a B plate of NGC288, which contains a standard photoelectric sequence (R. D. Cannon, personal communication), and has a limiting magnitude of 21.6±0.5 mag (±s.e.). Allowing for slight differences in exposure times ($\Delta m = -0.4$ mag), air masses ($\Delta m = +0.2$ mag) and seeing ($\Delta m = +0.2$ mag), we conclude that an appropriate average value for the limiting magnitude of AAT609 is 22.0±0.5 mag (±s.e.).

The corrected redshift of Fornax I (ref. 3) is $z=0.0050\pm 0.0003$ (±s.e.) and, if $H=55$ km s^{-1} Mpc^{-1}, this corresponds to a provisional distance of 27 Mpc. If the faintest of the suspect objects is characterised by an angular size of less than 2″ and an apparent B magnitude of 22.0 mag, then at this distance these parameters correspond to a linear dimension of less than 260 pc and an absolute B magnitude of about −10 mag. We suggest that these objects are unresolved globular clusters, although there are some difficulties associated with this hypothesis, which we discuss below.

The existence of globular clusters around a number of galaxies in the Virgo cluster[4,5] and, in particular, around the active elliptical M87 invites a comparison. Racine[4] has analysed the distribution of 2,000 globular clusters around M87 and more recently Hanes[5] has studied globulars associated with twenty galaxies (including M87) in the Virgo cluster. Hanes (personal communication) has revised Racine's conclusions regarding the luminosity function of the M87 globulars. He finds that there is no sharp cutoff at the bright end of the luminosity function, but rather a Gaussian tail with the brightest globular occurring at about $B=20.4$ mag The brightest of the Fornax objects lies at about 1 mag above the plate limit and has an apparent

B magnitude of 21.0±0.6 mag. The hypothesis that the upper limit to globular luminosities, around any given galaxy, is set by statistical fluctuations in random samples of various sizes, has been used by de Vaucouleurs[7] to substantiate an empirical linear correlation between the absolute magnitude of the first-ranked globular and the total absolute magnitude of the associated galaxy. We have used this correlation to obtain corrected distance moduli for M87 and NGC1399 of 30.4 mag and 30.8 mag respectively (allowing for 0.2 mag of extinction in both cases) and absolute magnitudes of the two galaxies of −21.1 mag and −20.2 mag for the corrected B magnitudes of 9.35 mag and 10.62 mag (ref. 8). The corresponding absolute B magnitudes of the brightest globulars in the two systems are at −10.0 mag and −9.8 mag. M87 has, however, been suspected to be abnormally rich in globular clusters (D. A. Hanes, personal communication) and a more detailed calculation will be presented later (de Vaucouleurs and J. A. D., unpublished). The difference in the distance moduli of M87 and NGC1399 is 0.4 mag, which corresponds to a distance ratio of 1.20. This may be compared with the ratio of the corrected redshifts[3] for the E cores of Fornax I and Virgo, which equals 1.45.

Fornax I lies within the Southern Supercluster which stretches from Cetus to Dorado[9]. The background density of faint galaxies is high and non-uniform, and the possibility that the clusters of faint objects around the three galaxies might result from chance alignments must be considered. This possibility is supported by the observation that there are few faint objects around NGC1404, even though it is a galaxy of similar luminosity to NGC1399. Although the system around NGC1399 is the most populous and extensive, the marked asymmetry of the distribution about the galaxy must raise doubts as to whether the two are physically related. Indeed, a Coma-like cluster of galaxies, removed to forty times the distance of the Coma cluster, would match the distribution both in angular extent and apparent magnitude of the 10th-ranked member. The number of objects around NGC1379 is sufficiently small to be accounted for as a local enhancement in the number of field objects, but in NGC1374 there is a system which gives the firm impression of a symmetrically placed distribution of globular clusters.

Smith and Weedman[10] have reported the discovery of globular clusters around NGC3311 in the Hydra I cluster of galaxies and, in common with others[4,6,7], have discussed their use as distance indicators leading to an evaluation of the Hubble constant. Fornax I is about 150° away from the supergalactic centre in Virgo and is well placed to reflect any apparent anisotropy in the Hubble flow which may arise from systematic motions within the local supercluster.

We thank Professor G. de Vaucouleurs, Dr M. G. Smith and the staff of ROE for helpful discussions, Dr H. A. Hanes for communicating his unpublished data and for many constructive suggestions, and also Dr R. D. Cannon for the loan of plate AAT609.

J. A. DAWE

Royal Observatory Edinburgh,
Edinburgh 9, UK

R. J. DICKENS

Royal Greenwich Observatory,
Herstmonceux Castle,
Hailsham, East Sussex, UK

1 Allen, C W., in *Astrophysical Quantities*, third ed. (Athlone, London, 1973)
2 Arp, H C , in *Galactic Structure* (University of Chicago Press, Chicago, 1965)
3 de Vaucouleurs, G., *Nature* (in the press)
4 Racine, R , *Jl R. astr. Soc. Can.*, 62, 367 (1968).
5 Racine, R., *Jl R. astr. Soc. Can.*, 64, 257 (1970).
6 Hanes, D A., thesis, Univ. Toronto (1975).
7 de Vaucouleurs, G , *Astrophys. J.*, 159, 435 (1970)
8 de Vaucouleurs, G., de Vaucouleurs, A., and Corwin, H G., in *Second Reference Catalogue of Bright Galaxies* (University of Texas Press, Austin, 1976)
9 de Vaucouleurs, G , *Vistas Astr.*, 2, 1584 (1956).
10 Smith, M. G., and Weedman, D. W., *Astrophys. J*, 205, 709 (1976).

Night-time reception of a solar radio event

ON March 30, 1976, at 2104–2106 UT, anomalous high-level radio bursts were recorded with a 48-channel polarised high resolution radio spectrograph[1] operating at a frequency range of 20.85–23.20 MHz. The observation was made at Kiiminki (65°05′N, 25°54′E), in the complete absence of interference of any kind and with the Sun and Jupiter well below the horizon. The antenna (two crossed log-periodic units in parallel) was directed towards North and elevated 30° for the purpose of recording scintillation spectra of the radio source Cassiopeia A. The spectrograph is designed primarily for observations of Jupiter[2] and its sensitivity for recording Cassiopeia A is marginal.

The anomalous bursts are shown in Fig. 1. The power level of the bursts exceeds the highest scintillation peaks of Cassiopeia A by 10–15 dB. The level is comparable to
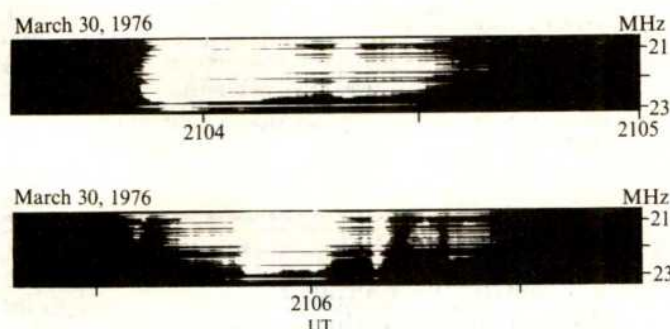

Fig. 1 Dynamic spectra of anomalous night-time radio bursts.

that of the strongest Jovian bursts observed close to the opposition of the planet. The colour[1,2] of the original record indicates that the right- and left-circular components of the emission had equal amplitudes.

At the same time, a strong solar radio event was in progress. From the radio spectra obtained at the CSIRO Solar Observatory, Culgoora, Australia; at the Harvard Radio Astronomy Station, Fort Davis, Texas (see Fig. 2); and at the Department of Astro-Geophysics, University of Colorado, Boulder, Colorado, it is inferred that the bursts are of solar type III. As can be seen from Fig. 2, the bursts start in the vicinity of 80 MHz. At the lowest recorded frequency (25 MHz) they match well with the bursts shown in Fig. 1.

Anomalous propagation of solar bursts have been observed previously[3]. In the present case the bursts must have been guided from the sunlit hemisphere over the polar regions. It may be significant that on the night of March 30, an auroral substorm and a strong blanketing sporadic-E (E_s) were observed (ref. 4 and T. Turunen, personal communication). A suitable duct may have existed between the ground and the E_s layer. This assumption is supported by

Fig. 2 The corresponding record of solar radio bursts, obtained at Fort Davis, Texas.
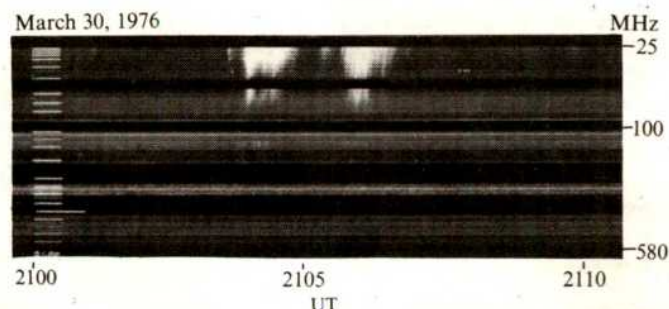
the low attenuation of the bursts. From their estimated flux density (A. Maxwell, personal communication) it is deduced that the total attenuation was less than 10 dB, probably just a few dB. The waves first penetrated the F2 layer, perhaps at a relatively high latitude, and then entered the duct beneath the E_s layer, provided the latter extended far enough towards the sunlit hemisphere.

The event was also recorded with a 20-MHz riometer at the Sodankylä Geophysical Observatory (67°25′N, 26°24′E). It appears as an outstanding peak signifying an increase in the noise level. It was not detected with the 27.6-MHz riometers at Sodankylä or at stations elsewhere in Finland (H. Ranta, personal communication) although strong solar emission occurred also at this frequency.

I thank Drs A. Maxwell, G. A. Dulk, and K. V. Sheridan for the Fort Davis, Boulder, and Culgoora solar records, and Mrs H. Ranta and Mr T. Turunen for the Sodankylä riometer and ionospheric data.

J. J. RIIHIMAA

Aarne Karjalainen Observatory,
University of Oulu,
Oulu, Finland

1 Riihimaa, J. J., *Aarne Karjalainen Observatory, Report No. 23* (1975).
2 Riihimaa, J. J., *Nature*, **255**, 210–211 (1975).
3 Smith, A. G., Carr, T. D., and Perkins, W. H., *Nature*, **183**, 597–598 (1959).
4 The Finnish Academy of Science and Letters, *Sodankylä Ionospheric Data and Geomagnetic Indices*, March 1976.

Absorption and the low velocity zone

IT is well known that elastic wave velocities are independent of frequency only for a non-dissipative medium[1]. In a real solid dispersion must accompany absorption. This effect has been emphasised by Randall[2] and Liu et al.[3]. The effect is small when the seismic quality factor Q is large or unimportant if only a small range of frequencies is being considered—that is, the spectra of P waves. Even in these cases, however, the measured velocities, or inferred elastic constants, are not the true elastic properties but lie between the high frequency and low frequency limits or the so-called 'unrelaxed' and 'relaxed' moduli. The magnitude of the effect depends on the nature of the absorption band and the value of Q. When comparing data taken over a wide frequency band the effect of absorption can be considerable especially considering the accuracy of present body-wave and free-oscillation data. Thus, Jeffreys[4] questions conclusions based on free oscillation data, particularly where they differ from body-wave results which are based on much shorter periods. Carpenter and Davies[5] and Davies[6] attempted to reconcile body-wave and surface-wave Earth models by allowing for physical dispersion using Futterman's[7] and Kolsky's[8] dispersion–absorption relationships. Jeffreys[4] used Lomnitz'[9] relationships. Liu et al. have shown that dispersion depends to first-order on absorption in the seismic frequency band and derived a linear superposition model that gives a Q that is independent of frequency. It can be shown that all of the above theories give equivalent absorption–dispersion relations for moderate absorption. Liu et al.[3] and Anderson et al.[10] have shown how to correct surface-wave and free-oscillation data for physical dispersion. Much of the support for the existence of an upper mantle low velocity zone has come from the inversion of normal mode data that have been uncorrected for physical dispersion due to absorption. In the light of these developments, we decided to re-examine the question of a shear-wave low velocity zone.

A starting model was constructed with monotonically increasing velocity and density in the upper 400 km of the mantle. The P and S velocities were chosen to satisfy the Jeffreys–Bullen[11] travel time tables to 30°. A smooth, Bullen[12] model A, density structure was used for this region. The model has a 3-km thick ocean layer, and an 18-km thick crustal layer.
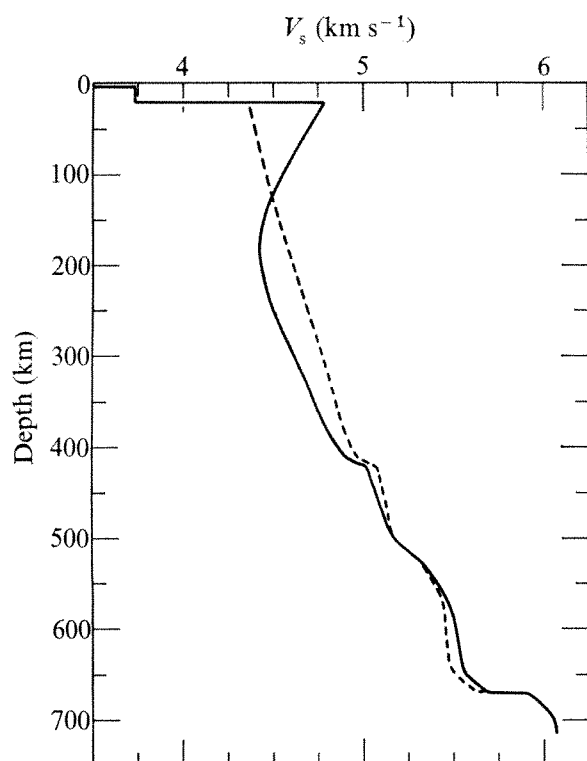
V_s (km s^{-1})

Fig 1 Shear velocity (V_s) as a function of depth for the starting model (dashed) and final model (solid).

These are average values for the Earth[13]. Below 400 km the parameters are the same as model C2 of Anderson and Hart[13].

Following Anderson et al.[10], the observed eigenperiods were corrected for the effect of absorption. A reference period of 1 s was selected to enable comparison with body-wave results. The corrected data included 66 fundamental spheroidal modes, 46 fundamental toroidal modes, 10 spheroidal overtones and 80 toroidal overtones. The spheroidal modes were selected for their sensitivity to shear velocity. The Q model MM8 of Anderson et al.[14] was used to compute the correction. This data set was then inverted, following the technique described by Jordan and Anderson[15]. The shortest smooth perturbation from the starting model which satisfies the data is found by this procedure.

The starting model does not provide a satisfactory fit to the normal mode data. The r.m.s. error is 0.7% whereas the r.m.s. error of the data is 0.09%. Only three modes are fit to within 1 s.d. of the data, and the errors for the shorter fundamental spheroidal modes (Rayleigh waves) are as large as 1.9%. The final model fits the data with an r.m.s. error of 0.07%; of the representative 78 modes used in the inversion, 62% fit the data to 1σ and 95% fit to 2σ.

The resulting upper mantle shear velocity (solid line) is shown in Fig. 1, together with the starting model. The new model has a pronounced low velocity zone with shear velocity decreasing from 4.77 km s^{-1} at the top of the mantle to a minimum of 4.42 km s^{-1} at 170 km. The average JB shear-wave travel time residual for the new model is +0.05 s over the distance range 30–95°. This can be compared with the 4–6-s discrepancy found from previous studies in which absorption was ignored[13,15].

We agree with Jeffreys[4], Davies[6], and Carpenter and Davies[5] that absorption must be taken into account when inverting surface-wave and free-oscillation data. Although some important revisions are required to Earth models determined from previous studies[13,15] we conclude that a low velocity zone is required by the observed data. There is no longer any 'baseline' discrepancy between body waves and normal modes. A more complete inversion which utilises the complete spheroidal mode data set will be reported elsewhere.

DON L. ANDERSON
ROBERT S. HART

Seismological Laboratory,
California Institute of Technology,
Pasadena, California 91125

1 Zener, C., *Elasticity and Anelasticity of Metals* (University of Chicago Press, Chicago, 1948).
2 Randall, M., *Physics of Earth and Planetary Interiors* (in the press).
3 Liu, H., Anderson, D. L., and Kanamori, H., *Geophys. J. R. astr. Soc.* (in the press).
4 Jeffreys, H., *Nature*, 215, 1365–1366 (1967).
5 Carpenter, E., and Davies, D., *Nature*, 212, 134 (1966).
6 Davies, D., *Geophys. J. R. astr. Soc.*, 13, 421–424 (1967).
7 Futterman, W., *J. geophys. Res.*, 67, 5279 (1962).
8 Kolsky, H., *Phil. Mag.*, 8, 693 (1957).
9 Lomnitz, C., *J. appl. Phys.*, 28, 201 (1957).
10 Anderson, D. L., Kanamori, H., Hart, R., and Liu, H., *Science* (in the press).
11 Jeffreys, H., and Bullen, K., *Seismological Tables* (Br. Ass. Adv. Sci., London, 1940).
12 Bullen, K., *The Earth's Density* (Wiley, New York, 1975).
13 Anderson, D. L., and Hart, R., *J. geophys. Res.*, 81, 1461–1475 (1976).
14 Anderson, D. L., Ben-Menahem, A., and Archambeau, C. B., *J. geophys. Res.*, 70, 1441 (1965).
15 Jordan, T., and Anderson, D. L., *Geophys. J. R. astr. Soc.*, 36, 411 (1974).

Effect of nearby supernova explosions on atmospheric ozone

WE have calculated the probable effects of a nearby supernova event on the ozone layer. We find that the effects are significant and long lasting, but are relatively rare at the location of the Earth in the Galaxy.

Ruderman[1] has suggested that cosmic-ray particles and γ rays produced by a nearby supernova could cause a substantial depletion of the Earth's ozone shield, with a consequently strong increase in the solar ultraviolet radiation incident on the surface of the Earth.

The suggested mechanism for ozone destruction is the formation by incident ionising radiation of large quantities of nitrogen oxide (NO_x) in the stratosphere. The NO_x then catalytically converts the ozone (and atomic oxygen) to molecular oxygen through the reaction sequence

$$NO + O_3 \rightarrow NO_2 + O_2 \text{ and } NO_2 + O \rightarrow NO + O_2$$
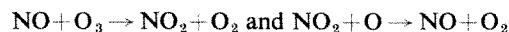
Ruderman has estimated the reduction in ozone column density that would result from a 10^{47}–10^{48}-erg burst of γ rays with a duration similar to that of a visible supernova event, that is, short compared with the residence time of NO_x in the stratosphere. He has also estimated the reduction that would result from an extended 'pulse' of cosmic rays with a total energy of ~ 10^{50}–10^{51} erg and a duration[2] of ~ 100 yr, the characteristic diffusion time over the distance from the supernova. The reductions obtained are quite dramatic; up to 90% in the case of the γ-ray burst.

We have applied more recent astronomical observations and somewhat different astrophysical concepts to the calculation of the intensity of ionising radiation expected from a supernova event. More refined models[3] of atmospheric chemistry and transport have been used to calculate the ozone depletion attributable to NO_x catalysis. We conclude that the ozone depletion, although smaller than that estimated by Ruderman, is still significant, and could, as a result of cosmic rays, extend over a very long period of time (10^3–10^4 yr); however, the probability of such an occurrence within the past 10^8 yr seems to be low.

In calculating the effect of γ rays, we first considered the possibility, as did Ruderman, that observed pulses of γ rays

in the 1–500-keV range[4] are attributable to extragalactic supernovae. This gives an upper limit on the energy content of a supernova γ-ray pulse of 10^{47}–10^{48} erg. Following Ruderman, we assume that all of the incident energy deposited at an altitude of 25–35 km. Our model enables us to follow the dynamics of the response of the atmosphere to a pulse, which we took to be a square wave of duration equal to the duration of the optical event (~ 100 d). Our calculations show that the occurrence of such an event 10 pc from Earth would result in peak ozone-column reduction of 35–65% for the 10^{47} and 10^{48} erg (upper limit) pulses, respectively. Such values are smaller than, but comparable to, Ruderman's estimates. As the e-folding recovery time of the stratosphere (the stratospheric residence time of NO_x) is ~ 3 yr, the time for the atmosphere to recover to 95% of normal would be ~ 10 yr.

Direct observations[5,6] of supernovae before, during, and after optical maximum give an upper limit to the emitted power in the 1–500-keV range of 10^{41} erg s^{-1}, which is compatible with the 10^{47}–10^{48} erg assumption discussed previously. Recent searches for smaller bursts[7] seem, however, to favour the hypothesis of galactic origin, reducing the estimated intrinsic burst energy by a factor $\sim 10^6$. Such an energy would be completely insignificant with respect to ozone reduction.

In previous calculations of the terrestrial effects of the cosmic rays produced by nearby supernova events, such as those by Ruderman[1] or by Laster[2], it has been assumed that the explosion accelerates a substantial fraction of the envelope to relativistic velocities, producing a 'pulse' of cosmic rays of energy comparable in magnitude to the total energy of the event (10^{50}–10^{51} erg). Recent spectroscopic observations suggest, however, that this is not the case[8], and tend to support models of the outburst which do not predict cosmic-ray 'pulses' of any significant energy[9]. It is known, however, that supernova remnants (SNR) expanding into the interstellar medium contain trapped cosmic-ray particles of high energy. The total energy content of SNRs can be deduced from the observed synchrotron radiation from trapped electrons. Our approach, then, is to assess the effects of envelopment of a planet in an expanding SNR containing trapped cosmic rays. The effects of such radiation would persist for the lifetime of the SNR ($\sim 10^3$–10^4 yr) and would be expected to be more biologically significant than a 10–10^2-yr effect of comparable magnitude.

To estimate the total energy contained in the form of trapped cosmic rays, W_{cr}, we have taken values of W_{cr} given by Ginsburg and Syrovatskii[10] for eight remnants with radii between 2 and 20 pc, and corrected them using more recently determined radii[11]. It seems that W_{cr} remains roughly constant over the lifetime of SNRs, at least in this size and age range, with a value of $W_{cr} \sim 2 \times 10^{49}$ erg. Assuming that the particle energy density decreases as r_0^{-3}, where r_0 is the radius of the SNR, and assuming also an isotropic particle velocity distribution with $|v| \sim c$, the intensity of cosmic radiation ($E \gtrsim 100$ MeV) at a point within the SNR is $\sim 10^{10}/r_0^3$ erg cm^{-2} sr^{-1} yr^{-1}, where r_0 is in parsec. We choose $E = 100$ MeV as a low-energy cutoff because cosmic-ray particles of energies greater than ~ 100 MeV are not deflected appreciably by the interplanetary magnetic field associated with the solar wind. The calculated intensity would then represent an increase over the current background level[12] of a factor of 10^2 for $r_0 \sim 10$ pc and about 10^3 for $r_0 \sim 5$ pc. Because of the lower initial energy we used, enhancements (an enhancement of 1 implies a doubling of the background level) are less than those calculated by Ruderman[1], but the duration of the exposure of a planet in the situation treated here would be about the lifetime of the SNR, $\sim 10^4$ yr, as opposed to the $\lesssim 10^2$ yr duration of an hypothetical 'pulse' of cosmic rays. In addition, significantly greater enhancements over shorter periods may be expected in the case treated here, because of the non-uniform distribution of particles within the SNR (for example, 'whisps' in the Crab and Cygnus or 'knots' in Cas A[13]).

The model calculations of atmospheric effects correspond to the steady state in this case, because the duration of the event is
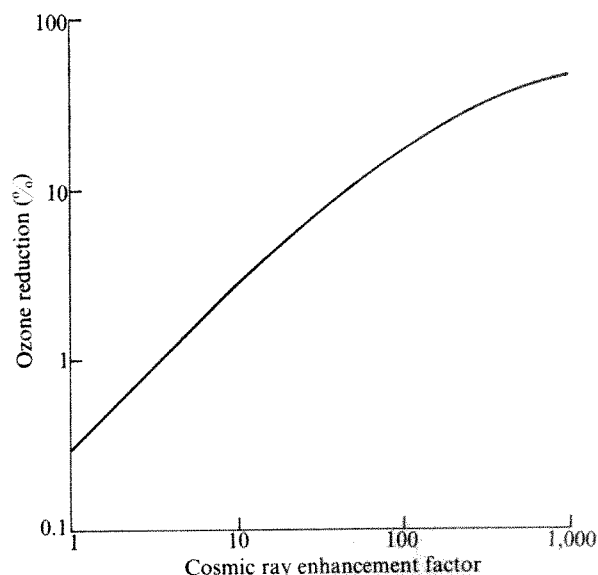
Fig. 1   Reduction (%) of ozone column density in the steady state as a function of the cosmic-ray enhancement factor.

very long compared to the stratospheric recovery time. The calculated global ozone reductions for various cosmic-ray enhancements up to a factor of 10^3 are shown in Fig. 1. It is seen that ozone reductions of ~ 20–50% could be expected from enhancements of 10^2–10^3. Although these reductions are smaller than estimated by Ruderman[7], they are still considerable, especially in the light of their probable long duration.

The waiting time, T_w, for a supernova outburst to occur within a volume $V_0 = (4\pi/3)r_0^3$ of the Galaxy is given by $T_w = T_{Gal}V_{Gal}/V_0$, where T_{Gal} is the galactic average time between supernovae in a total volume V_{Gal}. The distribution of supernovae in our Galaxy in space and time is still a matter of some uncertainty. Within 8 kpc of the galactic centre, SNRs seem to lie within a disk ~ 120 pc thick, and from 8 to 14 kpc the thickness is perhaps twice as large and the total volume density is $\sim 1/4$ as great as in the inner regions[11]. Using these observations and a mean time between supernova events of $T_{Gal} = 50$ yr, consistent with optical, SNR, and pulsar observations[14–17], we have found that at the location of the Earth (10 kpc from the galactic centre), waiting times for extended ($\sim 10^4$ yr duration) ozone depletions of 20–50% are between $\sim 2 \times 10^9$ and $\sim 2 \times 10^{10}$ yr. The probability that such an event occurred near (that is, ~ 5–10 pc from) the Earth within the past 10^8 yr is ~ 1–5%. Interestingly, the waiting times for such events to occur near a given planet at less than 8 kpc from the galactic centre are much shorter, $\sim 10^8$ and 10^9 yr, respectively.

It thus seems that the ozone depletion effect is significant and long lasting, and may affect the surface biology of the Earth or any Earth-like planet greatly, but is of low probability in our region of the Galaxy. Considering that cosmic and γ radiation does not penetrate to the surface in sufficient doses to affect life in a direct sense[1], and that ultraviolet light from a supernova at several pc is weaker than that from the Sun and of short duration (< 1 yr), the calculated ozone depletion would seem to be the major effect of a supernova on a nearby (~ 5–10 pc distant) Earth-like planet.

R. C. WHITTEN
J. CUZZI
W. J. BORUCKI
J. H. WOLFE

*Ames Research Center, NASA,
Moffett Field, California 94035*

1 Ruderman, M. A., *Science*, **184**, 1079 (1974).
2 Laster, H., *Science*, **160**, 1138 (1968).
3 Whitten, R. C., Borucki, W. J., and Turco, R. P., *Nature*, **257**, 38 (1975).
4 Klebsaedel, R. W., Strong, I. B., and Olson, R. A., *Astrophys. J. Lett.*, **182**, L85 (1973).
5 Canizares, C. R., Neighbours, J. E., and Matilsky, T., *Astrophys. J. Lett.*, **192**, L61 (1974).
6 Ulmer, M. P., Baity, W. A., Wheaton, W. A., and Peterson, L. E., *Astrophys. J.*, **193**, 535 (1974).
7 Bewick, A., Coe, M. T., Mills, J. S., and Quenby, J. J., *Nature*, **258**, 687 (1975).
8 Kirschner, R. P., Oke, J. B., Penston, M. V., and Searle, L., *Astrophys. J.*, **185**, 197 (1973).
9 Grassberg, E. K., Imshennik, V. S., and Nadyozhin, D. D., *Astrophys. Space Sci.*, **10**, 28 (1971).
10 Ginsburg, V. L., and Syrovatskii, S. I., *The Origin of Cosmic Rays*, 114 (Macmillan, New York, 1964).
11 Clark, D. H., and Caswell, J. L., *Mon. Not. R. astr. Soc.*, **174**, 267 (1976).
12 Meyer, P., *A. Rev. Astr. Astrophys.*, **7**, 1 (1969).
13 Bell, A. R., Gull, S. F., and Kenderdine, S., *Nature*, **257**, 463 (1975).
14 Tamman, G. A., *Astr. Astrophys.*, **8**, 458 (1970).
15 Milne, D. K., *Aust. J. Phys.*, **23**, 425 (1970).
16 Ilovaisky, S. A., and Lequeux, J., *Astr. Astrophys.*, **20**, 347 (1972).
17 Gunn, J. E., and Ostriker, J. P., *Astrophys. J.*, **160**, 979 (1970).

Heavy-ion damage in α Fe

WE summarise here some preliminary observations of damage in α Fe irradiated with low doses of heavy ions at room temperature. Although the damage produced by such irradiation has been studied in a number of pure metals and alloys using transmission electron microscopy[1], there have been comparatively few studies of materials of technological interest. Iron is of particular interest since it is a major constituent in many practically important alloys including ferritic steels. The use of heavy-ion irradiation (typically to doses $\lesssim 10^{13}$ ions cm^{-2} and energy $\lesssim 200$ keV) provides a convenient method of simulating the displacement cascades generated in metals undergoing fast neutron irradiation. In a majority of the metals and alloys studied the damage consists of a population of vacancy loops produced heterogeneously at the cascade sites by a collapse of the vacancy-rich cascade centres. Interstitial loops are not normally observed at the low doses commonly used, and this is thought to result from the close proximity of the damage region to the foil surface (the mean range of ions is typically ~ 100 Å) which acts as a dominant sink for the mobile interstitial point defects. There have been no previous reports of low-dose heavy-ion damage in α Fe and in other damage experiments[2] the vacancy component of the damage was not observed.

In this experiment, thin polycrystalline foils of α Fe were irradiated at room temperature with heavy ions of energy between 40 and 240 keV to doses $\simeq 5 \times 10^{12}$ ions cm^{-2}. The ions used were Fe$^+$, Ni$^+$, Ge$^+$, Kr$^+$, Xe$^+$ and W$^+$, with atomic weights ranging between 56 and 184. The iron used was the same material as that used in ref. 3 where the method of preparation of the specimens and an analysis of the iron purity is given. After irradiation, the foils were examined in a Siemens 102 or a JEM 100B electron microscope. A major experimental difficulty was the appreciable growth of oxide film, even though the specimens were stored under high vacuum (10^{-5} mmHg) after thinning. This, together with problems in producing foils with smooth surfaces and in completely eliminating objective astigmatism, meant that image quality was not as good as can be obtained in other materials. We estimate the visibility threshold for small loops in the best foils to be ~ 20 Å.

A detailed analysis of dynamical black–white images in foils irradiated with 80 keV W$^+$ ions has revealed the presence of two kinds of perfect loops with a/2 ⟨111⟩ or a ⟨100⟩ Burgers vectors. All the loops were of vacancy type. The results are consistent with the sequence proposed for the nucleation of interstitial loops[4]: aggregation initially on [110] followed by shear to a/2 ⟨111⟩ or a ⟨100⟩. The relative numbers of the two kinds of loops appear to depend sensitively on the grain orientation. A full analysis of these results will be reported later.

Information on cascade collapse in α Fe has been obtained from measurements of the efficiency with which various ions produce visible clusters. Results for different 80 keV ions are shown in Fig. 1, where the defect yield (defined as the fraction
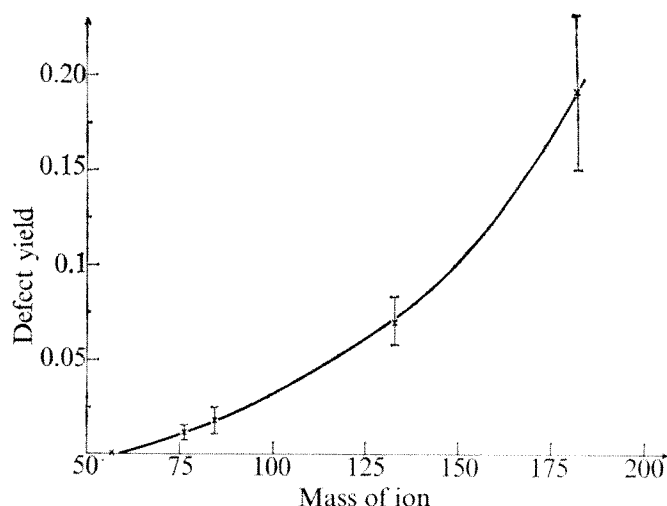
Fig. 1 Plot of defect yield against ion mass for grains close to ⟨110⟩ bombarded with 5×10^{12} 80-keV ions cm^{-2}.

of cascades which collapse to produce visible clusters) is plotted as a function of the ion mass. No correction has been made for loops slipping out of the foil, but care was taken that grains of similar orientation (in this case near ⟨110⟩) were selected for defect counting. The relative values of the defect yield should, therefore, be meaningful. Figure 1 illustrates two important results. Firstly, self-ions produce no visible damage; this was true for all ion energies in the range 40–240 keV. Secondly, the defect yield is a strong function of the mass of the incident ions, increasing from zero for Fe$^+$ ions to ~ 20% for W$^+$ ions, the heaviest ions investigated. The loops in α Fe irradiated with 80-keV W$^+$ ions ranged in size up to ~ 100 Å, but most were much smaller than this, ~ 90% falling in the size range 20–50 Å.

The absence of observable damage in self-ion irradiated specimens is consistent with results from neutron-irradiated α Fe[2] which showed that no visible damage is produced at doses $\lesssim 5 \times 10^{18}$ cm^{-2}. This suggests that cascades initiated by self-ions do not collapse, at least at room temperature and above. If this is so, it could have important consequences on the void swelling characteristics of α Fe, as have been discussed in ref. 5.

The explanation for the second observation concerning the sensitivity of the defect yield to ion mass cannot lie in differences in the total number of displacements produced, since, according to calculations[6], the ions used all produce approximately the same number of displacements, ~ 500 for 80 keV ions (the calculated numbers are shown in Table 1). What does alter, however, is both the spatial distribution of the damage and its depth below the foil surface. We do not consider that the latter affects the cascade collapse. Thus with increasing ion mass cascades become more compact, the vacancy supersaturation is higher and the cascade is hotter (in the sense that the density of energy deposited is increased). The magnitude of the effect is shown in Table 1. The quantities $\langle \Delta X^2{}_D \rangle$ and $\langle Y^2{}_D \rangle$ are related to the second moments of the energy distribution function and give measures of the cascade radius parallel and perpendicular to the incident ion direction respectively, while $\langle X_D \rangle$ gives a measure of the mean depth of the damage. These quantities have been computed from the WSS analytical theory[7]; θ_0 is an estimate of the maximum energy density in the cascade, calculated from ref. 8. In going from Fe$^+$ to W$^+$

Table 1 Calculations on collision cascades created by 80-keV ions

Ion	Fe	Ge	Kr	Xe	W
Atomic mass	56	76	84	131	184
No. of vacancies in cascade	507	513	518	535	548
$\langle X_D \rangle$ (Å)	193	169	156	105	95
$\langle \Delta X^2{}_D \rangle^{1/2}$ (Å)	120	102	96	64	59
$\langle Y^2{}_D \rangle^{1/2}$ (Å)	77	65	60	44	40
θ_0 (eV per atom)	0.21	0.37	0.42	1.26	1.81

ions the volume of the cascade decreases by about an order of magnitude, and the maximum density of deposited energy increases by about the same factor.

The work of English *et al.*[5] and English (unpublished), has highlighted the difficulty of cascade collapse in another b.c.c. metal, molybdenum. It was found that the defect yield produced by heavy-ion bombardment was dependent on both the irradiation temperature and on the mass of the incident ions. For 60-keV Mo[+] ions the defect density decreased at irradiation temperatures > 200 °C. This decrease cannot be explained on a vacancy-emission mechanism, where cascades are assumed to collapse athermally and then shrink by vacancy emission, nor by loop loss to the surface. The former mechanism had been successful in explaining the decrease in defect density in copper irradiated at high temperatures with 30-keV Cu[+] ions[10]. It seems that in self-ion irradiation of molybdenum the collapse process is itself temperature dependent. Further experiments with 60-keV Xe[+] and W[+] ions have, however, shown that the defect density in these cases remains unchanged at temperatures up to 425 °C, the highest irradiation temperature which could be achieved. It would seem, therefore, that in b.c.c. materials there is a sensitive balance between the collapse of vacancies in a cascade to a vacancy loop or their retaining an uncollapsed distribution and perhaps dispersing in the lattice.

Our results suggest that an important parameter in determining whether or not cascade collapse occurs in α Fe is the energy density or compactness of the cascades. The results from Mo irradiated with heavy ions at high temperatures provide further support for this concept and also suggest that an important and closely associated factor could be the rate at which vacancies diffuse out of the cascade centre into the surrounding crystal. There is considerable doubt regarding the temperature range for long range vacancy migration in α Fe and thus we are not able to separate this factor from the energy density factor in the present results. We plan to clarify this by irradiating iron specimens with ions of different masses at 77 K. The results from pure f.c.c. metals and alloys suggest that energy density in the cascade is a less critical factor governing cascade collapse than for b.c.c. metals. Nevertheless, there is a marked trend for higher cascade collapse efficiencies on going to heavy metals irradiated with self ions, Cu/Cu[+] → Ag/Ag[+] → Au/Au[+]. Moreover, Häussermann[11] has shown that in ion-irradiated copper the efficiency of cascade collapse to vacancy loops is higher for Au[+] ions that for Cu[+] ions, while Ruault *et al.*[12] have shown that in ion-irradiated gold the nature of the collapsed defects is also a function of ion mass. It seems, therefore, that energy density and cascade compactness are important parameters governing the development of damage structures in irradiated metals and their influence needs to be better understood.

C. A. ENGLISH
B. L. EYRE

Metallurgy Division,
AERE, Harwell,
Didcot, Oxfordshire, UK

M. L. JENKINS*

Department of Metallurgy
and Science of Materials,
University of Oxford,
Oxford, UK.

1 Eyre, B L , *J. Phys. F*, 3, 422–470 (1973).
2 Little, E. A , *Int. Metal Rev.*, 204, 25–60 (1976)
3 Eyre, B. L., and Bartlett, A. F., *Phil Mag.*, 12, 261–272 (1965).
4 Eyre, B. L., and Bullough, R., *Phil. Mag.*, 12, 31–39 (1965).
5 Bullough, R , Eyre, B L., and Krishan, K., *Proc. R. Soc.*, A346, 81–102 (1975).
6 Norgett, M. J , Robinson, M. J , and Torrens, I. M., *Harwell Report, AERE TP 494,* (1972)
7 Winterbon, K B , Sigmund, P., and Sanders, J. B , *Mat Fys. Medd. Dan. Vid. Selsk* , 37, 14–73 (1970)
8 Sigmund, P., *Appl. Phys. Lett.*, 25, 169–171 (1974)
9 English, C. A , Eyre, B L , Bartlett, A. F., and Wadley, H N. G., *Harwell Report, AERE R* 8083 (1976).
10 English, C. A., Eyre, B. L., and Summers, J., *Phil Mag.* (in the press).
11 Häussermann, F , *Phil Mag.*, 25, 537–599 (1972).
12 Ruault, M. O., Jouffrey, B., Chaumont, J , and Bernas, H., in *Proc. Conf. Fundamental Aspects of Radiation Damage in Metals, Gatlinburg, Tennessee,* 1 (edit. by Robinson, M T., and Young, F. W), 134–137 (USERDA, 1975)

Mechanical forces of electromagnetic origin

EXPERIMENTS have been reported which verified that at low frequencies a time-varying polarisation P in a dielectric, if taking place in a magnetic field H, results in a mechanical force of density $\dot{P} \times \mu_0 H$. This force is not predicted by the Minkowski energy momentum tensor, but the Abraham form of that tensor gives a force density in an homogeneous isotropic body of magnitude

$$[(\varepsilon_r \mu_r - 1)/c^2]\, \frac{\partial}{\partial_t}(E \times H)$$

usually called the 'Abraham Force'. In a dielectric this can be written $\dot{P} \times \mu_0 H + P \times \mu_0 \dot{H}$. It is not surprising that $\dot{P} \times \mu_0 H$ should give a mechanical force, since P corresponds to a movement of electric charge associated with polarisation, but it is difficult to understand why $P \times \mu_0 H$ should give a force.

To try to obtain experimental evidence of this effect, the same equipment was used as previously[1,2], namely a disk of barium titanate suspended as a torsional pendulum between the poles of a powerful electromagnet, the only change from the previous experiment was to hold P fixed but to vary H with time at the same frequency as the oscillation of the disk. Great difficulties were caused by the magnet system. A considerable coupling of energy to the pendulum resulted from mechanical vibrations in the magnet coils and there may also have been electromagnetic coupling because of small geometrical imperfections in the system. To avoid such resonance effects, we devised the following procedure.

If E and H have the same frequency, then, irrespective of their phase relationship, the time average of the Abraham Force is always zero. Thus if E and H are in time quadrature and if in fact $P \times \mu_0 H$ produces no mechanical force, there should be a time-averaged unidirectional force given by $\dot{P} \times \mu_0 H$. In the experiment, E and H were in time quadrature but at a frequency about 250 times the resonant frequency of the pendulum, thereby avoiding parasitic coupling from the magnet system. It was arranged that the direction of the electric field applied across the disk was reversed at the instants when the torsional pendulum reversed in its direction of motion, thereby ensuring that the applied force (if any) would always assist the oscillation of the pendulum. If no oscillation beyond the noise level resulted, the Abraham Force would be verified. In fact, a strong oscillation was observed which agreed closely (within 4%) with the amplitude predicted by the term $\dot{P} \times \mu_0 H$ acting alone.

Full details will be published elsewhere, but we conclude that neither the Minkowski tensor nor the Abraham tensor is consistent with the observed effects.

G. B. WALKER
G. WALKER

Electrical Engineering Department,
University of Alberta,
Edmonton, Canada

1 Walker, G B , and Lahoz, D. G., *Nature*, 253, 339–340 (1975).
2 Walker, G. B , Lahoz, D. G., and Walker, G , *Can. J Phys.*, 53, 2577–2586 (1975).

High resolution image of copper phthalocyanine

WE report here the successful high resolution of organic material by electron microscopy, a technique extremely sensitive to electron irradiation. As is well known, an inevitable factor that limits such high resolution microscopy is severe radiation damage, and, therefore, the electron beam must be reduced very much below the critical dose to prevent higher order details from fading. As a result, one cannot get high contrast
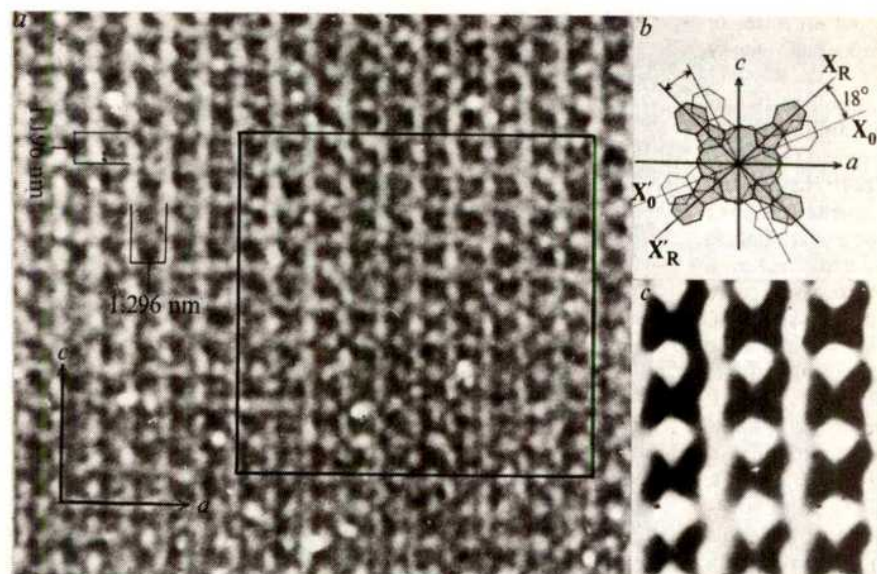
Fig. 1 Thin film for observation was prepared by vacuum deposition on a freshly cleaved KCl substrate, which had been preheated at 400 °C for 1 h. Film thickness was controlled carefully to be < 100 Å. We regard it as a perfect phase object. *a*, High resolution electron micrograph viewed along the *b* axis. Careful inspection reveals some crystal details. *b*, The relationship of the molecular orientation between the averaged imaged and the assumed isomorphic structure. $X_R X_R'$ indicates the molecular orientation of the retrieved image, $X_0 X_0'$ the assumed molecular orientation. *c*, The averaged image obtained by using the redundancy information from an enclosed region, by an optical method ($\times 1.24$).

by using the usual films, and we used an X-ray film (Ilford, Industrial G) with a high quantum efficiency. With this film, the exposure time did not exceed 5–7 s in conditions of minimal beam irradiation (Williams *et al.*[1]).

The specimen used was copper phthalocyanine, a widely used pigment. The thin film specimen for observation was grown epitaxially on a substrate of KCl. Preliminary experiments have found its crystal structure to be of the metastable α form[2], with crystal parameters: $a = 2.592$ nm, $b = 0.379$ nm, $c = 2.392$ nm, $\beta = 90.4°$, monoclinic, C2/c. Furthermore, its crystal structure was found by Ashida[3] to be closely isomorphic with platinum phthalocyanine (Robertson *et al.*[4]). The epitaxially grown crystallite has a special orientation relative to the substrate, namely, the *b* axis stands on the substrate surface at an oblique angle of ~ 32°, and planar molecules are stacked in parallel along the *b* axis. To observe the molecular image, therefore, the crystallite was tilted by a special holder with an oblique spacer. The observation was carried out in a JEM 100C.

The critical dose of copper phthalocyanine was found to be $1–2$ C cm^{-2} by measuring the complete fading of a diffraction spot (Reimer[5]). High resolution, however, requires many separate diffraction spots to construct the image, and therefore, as stressed above, the electron beam must be reduced below the critical dose. Although the electron dose was not directly measured, it was estimated as ≤ 0.1 C cm^{-2} from exposure meter readings.

Figure 1*a* shows a high resolution micrograph, taken at ~ 70 nm under focus. It can be seen that the micrograph is extremely noisy, a result of using minimal beam irradiation and X-ray film with a large grain. We used an objective aperture of 20 μm in diameter—through which spatial frequencies of up to 1.55 nm^{-1} could pass. Computer simulation would give the shape of the molecule under optimal focus condition (60–80 nm underfocus) but its shape would be poor[6]. We retrieved the molecular shape by using a lot of periodic unit cells, as described by McLachlan[7], and Fig. 1*c* represents a retrieved image from the region well defined in Fig. 1*a*, from which an optical diffraction pattern represented the considerable conservation of structure details with a slightly asymmetric intensity distribution from the original electron diffraction pattern. We can see the shape of the molecule clearly in Fig. 1*c*. Figure 1*b* represents the relationship between the averaged image and an assumed structure based on the close isomorphic relationship to platinum phthalocyanine, with a slight rotation of ~ 18°.

These pictures strongly suggest that a reconsideration of the crystal structure is in order. We are trying to improve our experiment by combining high resolution micrograph with

intensity data from electron diffraction patterns, as described by Unwin *et al.*[8].

We thank Mrs H. Murata for purification of crude copper phthalocyanine. Y. M. also thanks the SRC for support.

Y. Murata
J. R. Fryer
T. Baird

Chemistry Department,
The University of Glasgow,
Glasgow G12 8QQ, UK

1 Williams, R. C., and Fisher, H. W., *J. molec. Biol.*, **52**, 121 (1970).
2 Uyeda, N., Ashida, M., and Suito, E., *J. appl. Phys.*, 36, 1453 (1965).
3 Ashida, M., *Bull. Chem. Soc. Japan*, 39, 2632 (1966).
4 Robertson, J. M., and Woodward, I., *J. chem. Soc.*, 36 (1940).
5 Reimer, L., *Lab. Invest.*, 14, 1082 (1965).
6 Murata, Y., Fryer, J. R., and Baird, T., *J. Microsc.* (in the press).
7 McLachlan, D., *Proc. natn. Acad. Sci. U.S.A.*, 44, 948 (1958).
8 Unwin, P. N. T., and Henderson, R., *J. molec. Biol.*, 94, 425 (1975).

Hot lines in the Earth's mantle

THE mantle 'hot spot' or 'plume' hypothesis of Wilson[1] and Morgan[2] has been used to explain linear chains of volcanic islands and seamounts as well as aseismic ridges. These features of the ocean basins were presumably caused by the motion of lithospheric plates over mantle plumes, since it is assumed that unusually intense volcanism occurs at the Earth's surface above mantle plumes. Initially it was suggested that hot spots do not move relative to the lithosphere or to each other[1,2], but it was later proposed that some hot spots undergo a limited amount of motion relative to each other[3,4]. Here we hypothesise that some linear volcanic chains are the expression not of lithospheric motion over mantle hot spots but of 'hot lines' in the mantle, above which volcanism occurs intermittently.

The idea of mantle 'hot lines' is based on data collected on Easter Island (south-eastern Pacific) and along the tectonic-volcanic lineation running eastwards and westwards from the island. Wilson[1] and Morgan[2] originally suggested that Easter Island marks the site of one of the major 'hot spots' in the Earth's mantle. The island consists of volcanic rocks, mainly basalt, but with significant quantities of more acidic types, including mugearite, trachyte and rhyolite (ref. 5 and E.B., C.G.A.H., D. E. Fisher, J. Honnorez, J. G. Schilling, J. J. Stipp and M. Zentilli, to be published). The petrology and geochemistry of these rocks (to be published) show close similarities to rocks of other areas interpreted as expressions of mantle plumes, such as Iceland[6,7]. The Easter Island basalts are richer in alkalis, iron and titanium, and large-ion lithophile elements (such as Ba, Sr

and Zr) than ocean basalts. They are enriched in light rare earth elements, and have $^{87}Sr/^{86}Sr$, ^{207}Pb and $^{204}Pb/^{206}Pb/^{204}Pb$ ratios in a range higher than ocean ridge basalts but similar to Iceland basalts (to be published). Volcanism which gave rise to the island occurred less than 2 Myr before present, and mainly within the last 1 Myr (to be published). Thus, the petrology and age of Easter Island are consistent with a mantle plume origin.

If the Easter plume was active in the past, a trail of volcanoes should exist east of the island, along the direction of motion of the Nazca plate, and their age should increase with distance from the island. Indeed, a chain of volcanic islands and sea-mounts is observed along the so-called Easter Island Fracture Zone (to be published), including the islands of San Ambrosio and San Felix, and extending probably close to the Chile Trench (Fig. 1). The petrology of the islands of Sala y Gomez, San Felix, and San Ambrosio, and of a number of seamounts along the Easter Line, is consistent with a mantle plume origin (to be published). The ages of volcanism along the Easter Line, coupled with the spreading history of the Nazca plate, indicate, however, that the Easter volcanic chain cannot be explained by a fixed, single mantle plume.

On the basis of identification of magnetic anomalies, Herron[8] demonstrated that the East Pacific Ridge in the south-eastern Pacific became an active spreading centre about 9 Myr BP; before this, the fossil Galapagos Rise (Fig. 1) was an active spreading centre. Anderson and Sclater[9] reached similar conclusions using the age–depth relationship of Sclater *et al.*[10], even though their model of the spreading history of the Nazca plate differs in some respects from Herron's[8]. Either model would require the development of a gap of several hundred kilometres in the volcanic chain along the Easter line, if a fixed mantle plume under Easter Island is the cause of the volcanic chain (to be published). Our surveys and the topographic map of Mammerickx *et al.*[11] indicate that no such gap exists; therefore the Easter volcanic line cannot be caused by a single fixed mantle plume beneath Easter Island. Either a plume existed at the intersection of the Galapagos Rise and the Easter Line when the former was active, and moved to the Easter Island area when the Galapagos Rise became inactive, or two plumes coexisted, one in the Easter Island area, one at the Galapagos Rise (to be published).

The age of volcanism along the Easter line confirms that the volcanic chain could not have originated solely from the motion of the Nazca plate over the Easter Island hot spot. Sala y Gomez was active 1.7 Myr ago, so it could not have formed at the site of Easter Island unless one assumes unreasonably high spreading rates. San Felix and San Ambrosio, located more than 2,700 km from Easter Island, were formed not earlier than one or a few million years ago, and still show signs of activity (to be published).

We conclude that the activity which gave rise to the Easter volcanic line took place intermittently at several sites along the line (Fig. 1). This observation, coupled with the fact that the products of this activity are of the type ascribed elsewhere to mantle plumes, leads us to propose that a thermally anomalous linear zone exists in the mantle beneath the Easter line. This mantle hot line is the cause of the intense volcanism along the Easter line.

We find it unlikely that the Easter line is simply a fracture zone caused by thermal contraction of the lithosphere[12,13] and that the igneous activity along it is solely the passive result of the lithospheric fracture–fracture zones such as those from the East Pacific (Mendocino, Murray, Clipperton and so on) and from the Atlantic (such as Romanche, St Paul and Vema), presumably caused in part by lithospheric contraction, are characterised by little or no volcanism. The intensity of volcanism along the Easter line, and the fact that the magma source is probably deeper than the source of ocean ridge magmas, suggest that the Easter hot line reflects processes initiated beneath the lithospheric plates.

Based on theory and laboratory experiments, two types of mantle convective flow have been postulated by Richter and
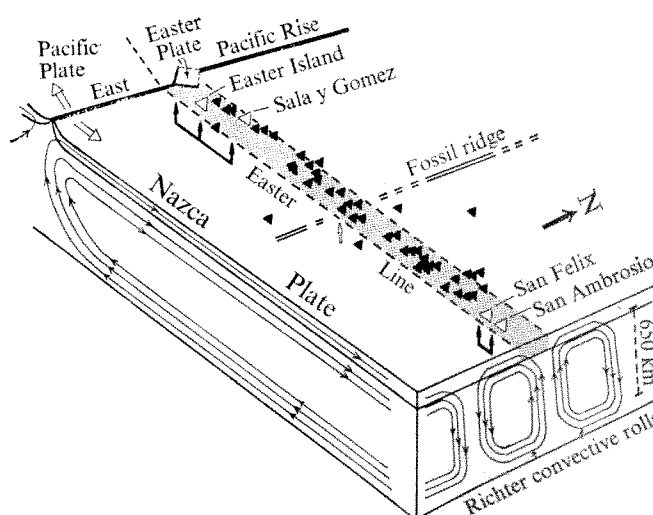
Fig. 1 Schematic illustration of the Easter hot line, showing the distribution of islands and of known seamounts in the area. Vertical full arrows indicate sites of young (< 2 Myr) volcanism; vertical open arrow indicates site of old (~10 Myr) volcanism at the intersection of the Easter line and the fossil ridge. The small Easter plate is described by Herron[16]. The Easter hot line is shown as due to Richter[14,15] mantle convective rolls; upwelling spouts[14,15] may actually be responsible for the line. △, Islands; ▲, seamounts.

Richter and Parson[14,15]. The first is a large scale flow, consisting of the plates themselves and of the return flow necessary to conserve mass, and the second a small scale flow, approaching a Rayleigh–Benard type convection, assumed to reach depths of about 650 km (where a major seismic discontinuity exists), with a similar horizontal scale[14,15]. The small scale convection rolls tend to become aligned with their axes parallel to the direction of motion of the plate—in times of 20–50 Myr for fast moving plates (10 cm yr^{-1}) but as much as several hundred million years for slow moving plates (2 cm yr^{-1}) (refs 14 and 15).

We suggest that the Easter hot line is the expression of mantle activity occurring along the rising limbs of adjacent Richter[14,15] convective rolls, with their axes parallel to the direction of motion of the Nazca plate. The required condition of relatively high plate velocity is probably fulfilled or at least approached in the Nazca plate, if we assume that the velocity of plates relative to the sublithospheric mantle is related to the measured spreading rates (~ 10 cm yr^{-1}) (refs 8 and 9). Making the same assumption, the direction of the Easter hot line is also the correct direction for a Richter convective roll, as it is almost exactly parallel to the direction of spreading. On the other hand, at lower plate velocities a pattern of mantle flow develops which is characterised by both upwelling and downwelling spouts, also aligned parallel to the direction of plate motion[12].

If the horizontal scale of each cell in the transverse convective rolls is about 650 km (refs 14 and 15), the distance between upwelling spouts would be roughly twice that distance. The San Felix–San Ambrosio 'hot' area is roughly 2,800 km from Easter Island or 2,600 km from a middle point between Easter Island and Sala y Gomez, and the fossil 'hot spot' at the Galapagos Rise was roughly halfway between the Easter–Sala y Gomez and the San Felix–San Ambrosio hot area. Thus, the distance between hot areas along the Easter line (roughly 1,300 km) is consistent with Richter and Parson's model.

We conclude that the Easter hot line is the result of mantle convection rolls or spouts aligned parallel to the direction of plate motion. The small Easter Plate[16] may be the result of complex interaction of two patterns of mantle flow at the intersection of the Easter line and the East Pacific Rise. It is possible that a hot line exists also west of the East Pacific Rise, extending into the Pacific Plate, along a chain of islands and seamounts up to Pitcairn Island. Pitcairn is < 1 Myr old[17].

Hot lines differ from 'normal' fracture zones in that the latter are primarily the result of lithospheric contraction[12,13], with the sublithospheric mantle having a passive role and consequently with little or no mantle-derived volcanism. The 'hot lines' are instead caused by activity of the sublithospheric mantle, as in Richter and Parson's model, and are characterised by volcanism of a type previously ascribed to mantle plumes. It is possible that some oceanic aseismic ridges are caused by mantle 'hot lines' rather than by motion of plates over hot spots.

This research was supported by the US Office of Naval Research and by NSF.

E. BONATTI

Lamont–Doherty Geological Observatory,
Columbia University,
Palisades, New York 10964

C. G. A. HARRISON
Rosenstiel School of Marine and Atmospheric Science,
University of Miami,
Miami, Florida 33149

1 Wilson, J. T., *Phil. Trans. R. Soc.*, **258**, 145–167 (1965); *Tectonophysics*, **19**, 149–164 (1973).
2 Morgan, W. J., *Nature*, **230**, 42–43 (1971).
3 Molnar, P., and Atwater, T., *Nature*, **246**, 288–291 (1973).
4 Molnar, P., and Francheteau, J., *Geophys. J. R. astr. Soc.*, **43**, 763–774 (1975).
5 Baker, P. E., Buckley, F., and Holland, J. G., *Contrib. Miner. Petrol.*, **44**, 85–100 (1974).
6 Schilling, J. G., *Nature*, **242**, 565–571 (1973).
7 Hart, S. R., Schilling, J. G., and Powell, J. L., *Nature*, **246**, 104–107 (1973).
8 Herron, E. M., *Bull. geol. Soc. Am.*, **83**, 1671–1692 (1972).
9 Anderson, R. N., and Sclater, J. G., *Earth planet. Sci. Lett.*, **14**, 433–441 (1972).
10 Sclater, J. G., Anderson, R. N., and Bell, M. L., *J. geophys. Res.*, **76**, 7888–7915 (1971).
11 Mammerickx, J., Anderson, R. N., Menard, H. W., and Smith, S. M., *Bull. geol. Soc. Am.*, **86**, 111–118 (1975).
12 Collette, B. J., *Nature*, **251**, 299–300 (1974).
13 Turcotte, D. L., *J. geophys. Res.*, **70**, 2573–2577 (1974).
14 Richter, F. M., **78**, *J. geophys. Res.*, 8735–8745 (1973).
15 Richter, F. M., and Parson, B., *J. geophys. Res.*, **80**, 2529–2541 (1975).
16 Herron, E. M., *Nature*, **240**, 35–37 (1972).
17 Duncan, R. A., McDougall, I., Carter, R. M., and Coombs, D. S., *Nature*, **251**, 679–682 (1974).

Acoustic response to chemical stimuli in ground crickets

IN several species of ground crickets (Orthoptera, Gryllidae, Nemobiinae), chemical communication may be used in addition to acoustic communication in bringing males and females together. In laboratory experiments, a high proportion of males of three *Allonemobius* species and one *Pictonemobius* species showed rapid antennation and produced calling songs shortly after exposure to paper from the cages of female crickets. This is the first report of an acoustic response to chemical stimuli in singing Orthoptera.

Populations of *A. fasciatus*, *A. allardi*, *A. tinnulus* and *Eunemobius carolinus* were studied where they occur in sympatry in Suffolk County, Long Island, New York, in 1972, 1973 and 1974. *A. fasciatus* and *A. allardi* were also studied in Tompkins County, New York in 1974. In addition, sympatric populations of *A. fasciatus* and *P. ambitiosus* were studied in Gainesville, Florida, in 1975. Several *A. sparsalsus* males were collected in Florida (by T. J. Walker) and then studied in New York in 1974. The male crickets used in experiments were field captured, except for 33.3% of the *A. tinnulus* males and 18.9% of the *A. fasciatus* males, which were laboratory reared.

Adult males were kept individually in small cages for 1 week before testing. The cage of each male was carried from a rearing area to an experimental area where the cricket was placed in a plastic box (35.5 × 28 × 15 cm) containing one of the following stimuli: (1) control, clean paper towels; (2) paper conditioned in the rearing cages of female crickets (usually 12 virgin females); (3) paper conditioned by males (groups of 8–12 adult male crickets); (4) live females (usually 12 virgin females). All experiments were carried out under laboratory light during the 'on' periods of the rearing room light cycle, at temperatures between 22 and 28 °C. In experiments with paper, a stopwatch

Fig. 1 Proportions of *A. fasciatus* (*a*), *A. allardi* (*b*) and *A. tinnulus* males (*c*) that sang in tests with control paper (C), female-conditioned paper (♀P) and live females (♀♀). For each species, a significantly greater proportion sang on female-conditioned paper than on control paper. A significantly greater proportion of *A. allardi* males sang in the presence of live females than on female-conditioned paper. Numbers above bars refer to number of males tested.

was started when the male was dropped into the box; in experiments with live females, timing was begun when the male made antennal contact with a female. The watch was stopped at 5 min, or when the male began to stridulate. A stridulatory response within 5 min was noted as a positive response. Each male was tested only once with a given stimulus; for stimuli other than controls, males were tested only once every 24 h.

In addition to the main series of experiments, *A. fasciatus* males were tested with paper from the cages of early instar and older male *A. fasciatus* juveniles (that is, juveniles without visible ovipositors).

Many of the males gave a vigorous positive response in tests with paper conditioned by females (Figs 1 and 2). These males walked in the test box for several seconds after being placed on the paper, slowly antennating and palpating the substrate. Often, the cricket showed an obvious increase in the rate of antennation and made frequent erratic changes in body orientation, as though searching. Soon, usually within seconds, the male raised his tegmina (forewings) and stridulated. Some males stood still and sang, others slowly walked while singing.

A wide range of behaviour was shown by males that did not respond with song to conditioned paper. Some were completely inactive in the test box, others made the transition from ex-

Fig. 2 Proportions of *P. ambitiosus* (*a*), *E. carolinus* (*b*) and *A. sparsalsus* males (*c*) that sang in tests with control paper (C), female-conditioned paper (♀P) and live females (♀♀). A significantly greater proportion of *P. ambitiosus* males sang on female-conditioned paper than on control paper. Numbers above bars refer to number of males tested.
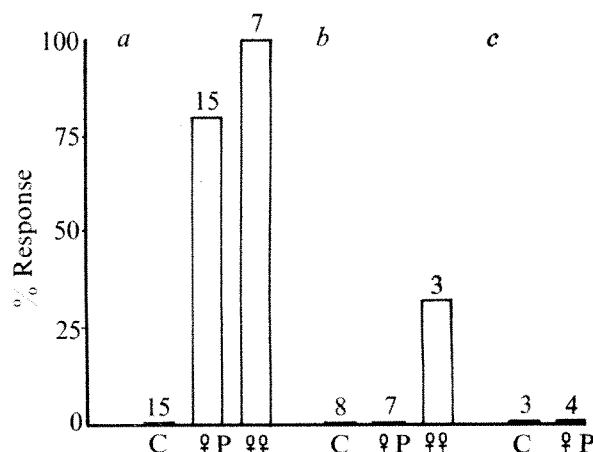
ploratory walking to the quick searching behaviour, but produced no song within 5 min.

The numbers above the bars in Figs 1, 2 and 3 are the numbers of males tested with each stimulus; the height of the bars shows the proportion of males that sang within 5 min. For *A. fasciatus*, *A. allardi*, *A. tinnulus* and *P. ambitiosus*, the proportion of males that sang in response to paper conditioned by conspecific females was significantly greater than the proportion that sang on the control paper ($P < 0.05$ in 2×2 *G* tests of independence[1]; Figs 1 and 2). In the absence of visual, acoustic, or tactile cues, males of the four species sang in response to some factor, presumably chemical stimuli left on the substrate by females. None of the *Eunemobius carolinus* and *A. sparsalsus* males gave a positive response in the presence of conspecific female paper.

A. fasciatus, *A. allardi*, and *A. tinnulus* males showed a significantly greater response to conspecific female paper than to paper conditioned by conspecific males (Fig. 3). Furthermore, *A. fasciatus* males showed a greater response to paper conditioned by adult females than to paper conditioned by *A. fasciatus* juveniles (Fig. 3). These results suggest a sexual role for chemical communication in these species. Ground crickets of these genera produce quiet, high pitched songs relative to those of field crickets (subfamily Gryllinae), or even other nemobiine genera (for example, *Eunemobius*)[2]. The softer, higher pitched songs are not likely to carry far to attract females at a distance. Perhaps the song-releasing stimuli left on the substrate by females are pheromones which indicate areas where singing males are most likely to attract females. The stimuli may be important in species isolation; I have experimental evidence of specificity at the species level[2].
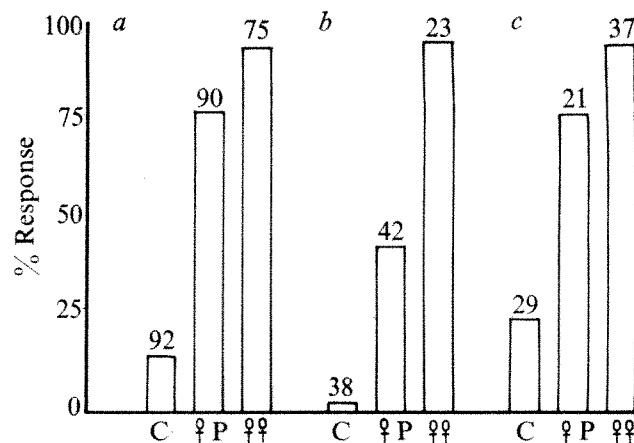

Fig. 3 Proportions of *A. fasciatus* (*a*), *A. allardi* (*b*) and *A. tinnulus* males (*c*) that sang in tests with female-conditioned paper (♀P) male-conditioned paper (♂P); and juvenile-conditioned paper (JP) for *A. fasciatus*. For each species a significantly greater proportion sang on female-conditioned paper than on male-conditioned paper. A greater proportion of *A. fasciatus* males sang on female-conditioned paper than on juvenile-conditioned paper. Numbers above bars refer to number of males tested.
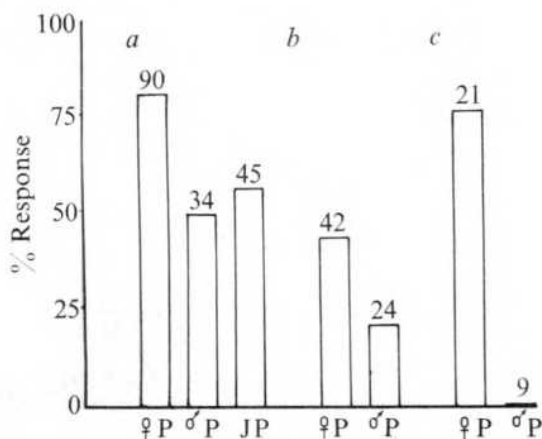
Previous studies have described pheromone-like communication in crickets[3,4] and chemically-induced stridulation has been reported for the Douglas fir beetle[5]. The data reported here, however, demonstrate song-releasing chemical stimuli in insects well known for their acoustic communication. Further studies may reveal something about the evolutionary origins of cricket song. In some cockroach species (for example, *Blattella germanica*) males raise their wings and expose gustatorial glands (which females feed on) in response to antennal contact with a female pheromone[6,7]. This is of interest here since crickets are believed to have a cockroach-like ancestor[8].

At present, the origin and nature of the song-releasing stimuli are unknown. Initial attempts to isolate stimuli from the faecal pellets of females have resulted in ambiguous bioassays.

I thank Drs T. J. Walker and R. Hoy for comments; the

Section of Neurobiology and Behaviour, Cornell University, for laboratory space. Early experiments were supported by the NSF; recent work was supported by the NIH.

R. PAUL

Department of Entomology,
University of Florida,
Gainesville, Florida 32611

1 Sokal, R., and Rohlf, J., *Biometry*, (Freeman, San Francisco, 1969).
2 Paul, R., Thesis, State Univ. New York, (Copies available, Univ. Microfilms, Ann Arbor, Michigan, 1975).
3 Sexton, O., and Hess, E., *Biol. Bull.*, 134, 490–502 (1968).
4 Otte, D., and Cade, W., *Anim. Behav.*, 24, 1–6 (1976).
5 Rudinsky, J. A., and Michael, R. R., *Science*, 175, 1386–1390 (1972).
6 Roth, L., and Willis, E., *Am. Midl. Nat.*, 47, 66–129 (1952).
7 Nishida, R., Fukami, H., and Ishii, S., *Experientia*, 30, 978–979 (1974).
8 Sharov, A. G., *Acad. Sci. U.S.S.R., Trans. Instit. Paleont.*, 18, (1968).

Changeable coloration of cornea in the fish *Hexagrammos octogrammus*

ALTHOUGH fish are considered 'lower' vertebrates, the visual apparatus of some species has a high degree of adaptability unparallelled in other vertebrates. For example, members of the Hexagrammidae can change their cornea rapidly from colourless in the dark, to deep red in bright light[1,2]. We have studied this ability in *Hexagrammos octogrammus* Pallas, a common shallow-water fish of the Japan Sea, and found that it is caused by the effects of illumination on the distribution of coloured cytoplasm in the corneal chromatophores.

Specimens of *H. octogrammus* caught by night or kept in the dark have a cornea as transparent and colourless as that of most fish. But specimens collected in their natural environment by day, or kept in an illuminated aquarium, have a bright orange or deep red cornea. If a light-adapted fish is placed in the dark, its cornea gradually pales, becom-

Fig. 1 Absorption spectra measured in pupil zone of corneas from *H. octogrammus* in different adaptation states: *a*, in full darkness; *c*, intermediate (yellow) coloration observed after about 30 min in sunlight after full dark adaptation; *d*, full coloration after 2 h in sunlight; *b*, yellow cornea of *Pleurogrammus monopterigius* in full colour; *e*, control.
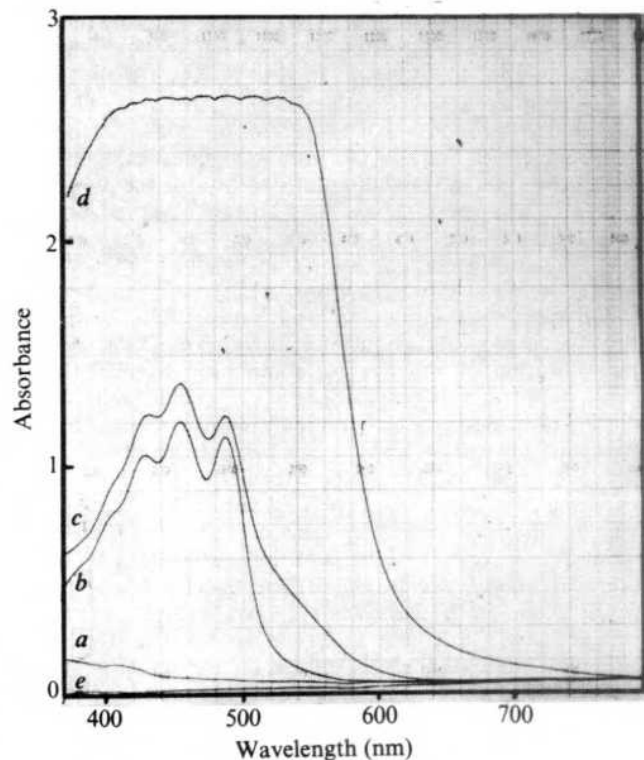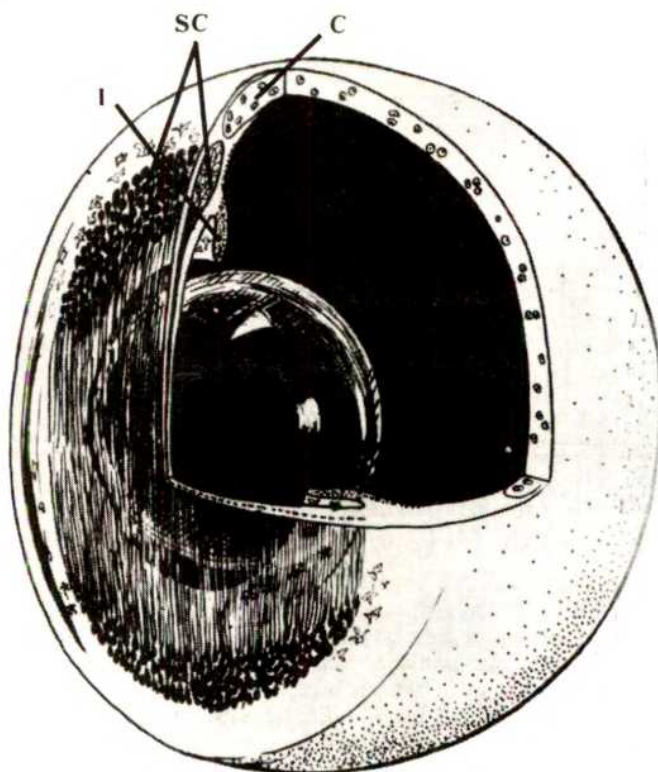
Fig. 2 Reconstruction of anterior part of the eye in *H. octogrammus*. SC, Chromatophores; C, chondreal sclera; I, iris.

ing yellow and then colourless. Similarly, in bright illumination, a colourless cornea becomes yellow and then red, even at night. The speed of colour change depends on the intensity of illumination, and when a fish is transferred from direct sunlight to darkness, complete decoloration usually takes 1.5–2 h. If the fish is kept continuously in light intensities of about 10 lx, its cornea becomes yellow without reddening.

The cornea of *H. octogrammus* has surprising absorptive properties as a colour filter. Its transmittance in the pupil zone in the light-adapted state does not exceed 0.3% for wavelengths of 400–500 nm. It absorbs more than 10% of the light up to a wavelength of 580 nm (Fig. 1). The most saturated coloration reported so far for a fish cornea has been that of the South American *Astronotus* (Cichlidae)[3]— yet in this species the minimum transmittance is not less than 4%, absorption above 10% being observed only for wavelengths shorter than 510 nm.

This changing corneal coloration can be compared with retinomotor and pupil reactions, which regulate the quantity of light reaching the receptor layer. The most likely function of the yellow and orange coloration is to improve the resolution of the eye by eliminating harmful loss of contrast of retinal image due to (1) a blue veil of light scattered on intraocular and extraocular media, and (2) chromatic aberration[4], which can be quite significant in the fish eye[5]. Thus shallow-water fishes like *H. octogrammus* searching for food on a well-lit substratum can benefit from a coloured cornea in spite of the loss of sensitivity because it is only at low light intensities that preretinal filters become a disadvantage. The exceptional saturation of the filters in *H. octogrammus* is tolerated because the fish can, so to speak, remove its sunglasses when it gets dark.

In most animals there is diffuse coloration of some part of the eye—in tree shrew, ground squirrels and diurnal snake the whole lens is yellow[4], in others, as in some squid[6]. a distinct layer of lens or of the cornea, as in the perch and pike, is coloured. In *H. octogrammus* the cornea is coloured

by specialised corneal chromatophores, which differ considerably in structure from ordinary dermal chromatophores.

Chromatophore cell bodies comprise two compact sickle-shaped masses near the upper and lower edges of the cornea, outside the pupil zone. In a vertical section of the cornea, cell masses are surrounded by fibrous tissue. Each chromatophore (50–70 μm in diameter) is pear shaped and gradually tapers into a single, flattened process. All processes are oriented vertically (in the eye of a living fish) and extend downwards or upwards, according to the position of the cells in the upper or lower cell mass, and overlap the pupil zone (Fig. 2).

Corneal coloration is altered by redistribution of coloured cytoplasm between cell bodies (Fig. 3) and their processes. Direct observation of the living fish under a dissecting microscope showed that in fish kept in bright light the processes are filled with cytoplasm, but in the dark they become empty and invisible.

There are two kinds of corneal chromatophore—deep red and yellow—resembling dermal erythrophores and xanthophores in their coloration. When a dark-adapted fish is exposed to bright light, the processes of the yellow chromatophores fill up first, giving a transient yellow colour to the cornea (Fig. 1c). The corneas of *Pleurogrammus monopterigius* and *H. stelleri* (both Hexagrammidae) have only yellow chromatophores, the colour of their corneas in the light-adapted state being similar to the intermediate coloration of that of *H. octogrammus* and the permanent colour of the cornea of pike and perch[7]. The absorption bands at 427, 455 and 485 nm reveal the carotenoid nature of the yellow pigmentation. The coloration of red cornea of *H. octogrammus* in light-adapted fish corresponds to absorption in the red corneal chromatophores, their absorption spectrum consisting of a single broad band (the half-band about 400–550 nm) with a single flattened maximum at

Fig. 3 Ultrastructure of the chromatophore from *H. octogrammus* cornea. Numerous lipid granules (LG), arborised invaginations of cell membrane (I), collagen fibres (C), pinocytosis vesicles (PV), mitochondria (M) and Golgi apparatus (GA) are visible. (× 29,750).
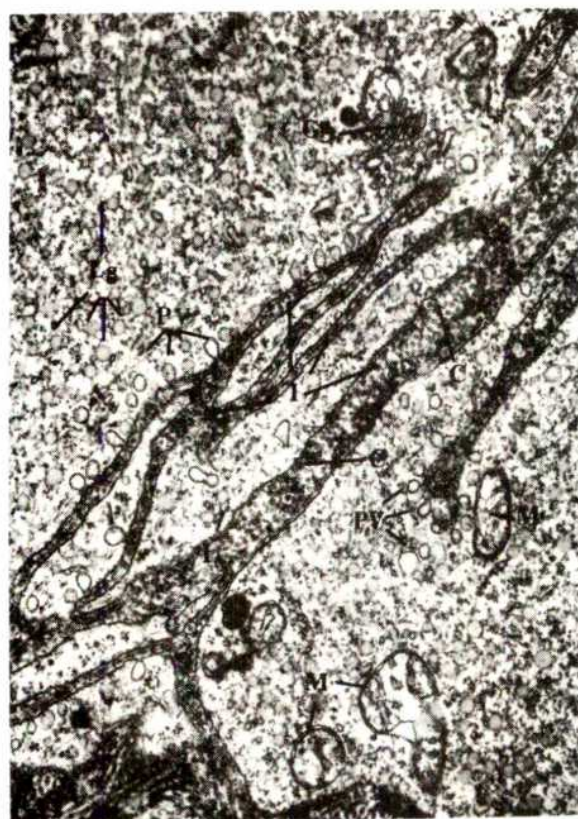
about 480 nm. The difference between the absorption spectra of the two types of corneal chromatophore, shown by microspectrophotometry, cannot be explained as resulting from differences in concentration of the same coloured substance.

The corneal chromatophores differ from other (dermal) chromatophores in having only one process, and having, in the cytoplasm of the cell body and process, homogeneous round granules of intermediate electron density (100–120 nm in diameter), which might be lipid droplets in which carotenoids are dissolved (Fig. 3). Both cell bodies and processes contain many microtubules (25 nm in diameter), lining the axis. No lipid granules have been found in the processes of dark-adapted fish, although microtubules and numerous pinocytosis vesicles are sometimes present. There are many invaginations of the cell membrane, up to 100 nm wide and several μm long, with pinocytosis vesicles attached to them. Unlike ordinary dermal chromatophores—erythrophores and xanthophores—the corneal chromatophores contain no pterinosomes[8].

Microtubules are directly concerned with dispersion and aggregation of pigment granules in melanophores[9,10], and antimitotic agents (colchicine and vinblastine) which destroy the normal function of microtubules may block the movement of pigment granules[11]. After injection of colchicine into the ocular bulb of *H. octogrammus* the response of cornea to a change of illumination is slower and weaker.

Changes of the corneal coloration caused by specific corneal chromatophores are not restricted to members of the Hexagrammidae. They have been observed, but to a lesser degree in several species of other families, such as the Blennidae, Cottidae and Tetraodontidae.

O. Y. ORLOV

Institute for Problems
 of Information Transmission,
Academy of Sciences of the USSR,
Moscow, USSR

A. G. GAMBURTZEVA

Institute of Marine Biology,
Far East Scientific Centre,
Academy of Sciences of the USSR,
Vladivostok, USSR

1 Orlov, O. Y., *Priroda*, **8**, 35 (1974) (in Russian).
2 Orlov, O. Y., Gamburtzeva, A. G., and Alexandrova, M. A., *Priroda*, **4**, 111 (1974) (in Russian).
3 Muntz, W. R. A., *Vision Res.*, **13**, 2235 (1973).
4 Walls, G. L., *The Vertebrate Eye and its Adaptive Radiation*, (Cranbrook Institute of Sciences, Michigan, 1942).
5 Sivak, J. G., *Vision Res.*, **14**, 209 (1974).
6 Denton, E. G., and Warren, E. J., *Nature*, **219**, 400 (1968).
7 Bridges, C. D. B., *Vision Res.*, **9**, 435 (1969).
8 Bagnara, J. T., Taylor, J. D., and Hadley, M. C., *J. Cell Biol.*, **38**, 67 (1968).
9 Bikle, D., Tilney, L. D., and Porter, K. R., *Protoplasma*, **61**, 322 (1966).
10 Green, L., *Proc. natn. Acad. Sci. U.S.A.*, **59**, 1179 (1966).
11 Murphy, D. B., and Tilney, L. G., *J. Cell Biol.*, **61**, 757 (1974).

Aversive behaviour of crown-of-thorns starfish to coral evoked by food-related chemicals

WHEN *Acanthaster planci*, the coral-predating crown-of-thorns starfish, encounters living coral or when it is presented with coral extracts, it rears the arms near to the stimulus source aborally and retracts the tube feet of these arms into the ambulacral groove[1,2]. This aversive behaviour has generally been attributed to the effect of discharge of coral nematocysts[1-3], or to toxins released from nematocysts[4]. But arm rearing may be evoked before contact is made with corals, and both "withdrawal" responses are produced by non-coral food or food extracts which stimulate *A. planci* to feed (R.J.M., unpublished). This suggests that it is not nematocysts or their toxins which are responsible, but rather chemicals from coral tissue.

The experiments described here show that this is the case: the function of the aversive response seems to be protection of the starfish against nematocyst discharge on contact with coral.

The role of nematocyst toxins was investigated by comparison of the effects on *A. planci* of extracts of whole coral tissue and homogenised isolated nematocysts from three coral species, including *Millepora dichotoma* ('fire coral'), which is only rarely attacked by *A. planci*. To validate comparisons between the extracts of each species of coral, and show up any differences in activity, it was necessary to equalise concentrations of non-stimulatory components of the extracts. Collins[5] found that only low molecular-weight components of coral extract, apparently amino acids, elicited arm withdrawal. It was assumed that soluble protein represented the bulk of non-stimulatory material in both nematocysts and tissue cells, and so the protein concentrations of the extracts were equalised—by dilution of the more concentrated extracts—after estimation by the method of Lowry *et al.*[6].

Table 1 summarises the effects of dripping extracts near the tips of arms of moving starfish. It shows that the homogenised nematocyst preparation of only one coral, *Millepora dichotoma*, evoked stronger responses than those produced by the corresponding whole coral extract, and even then the responses evoked were weaker than those produced by the whole tissue extract of *Acropora multicaulis*, in spite of the far greater protein concentration of the *Millepora* extract.

Moreover, the high efficiency of isolation of nematocysts of *Millepora dichotoma* (Fig. 1) made possible an estimate of the proportion of total tissue volume in intact *Millepora* occupied by nematocysts: it was approximately 2%. Thus the contribution made by nematocyst contents to the arm-rearing and tube-foot retraction activity of a whole coral extract seems to be negligible.

We even found that intact isolated nematocysts had no observable effect on the tube feet of *A. planci*—yet accidental splashes of nematocyst suspensions on to our skin produced typical stinging sensations.

These unexpected results suggested that nematocyst contents had leaked into the whole coral extracts during preparation, although most nematocysts remain undischarged. This objection was negated by testing extracts of mesenterial filaments, which contain greater concentrations of nematocysts than other coral tissues (Fig. 2) and may be removed intact from corals with large polyps. Even so, as Table 1 shows, the responses elicited by the mesenterial

Fig. Photomicrograph of nematocysts isolated from 'fire coral, *Millepora dichotoma*, and used in the preparation of nematocyst extracts. The darker objects are fragments of coral skeleton. For further explanation, see text.
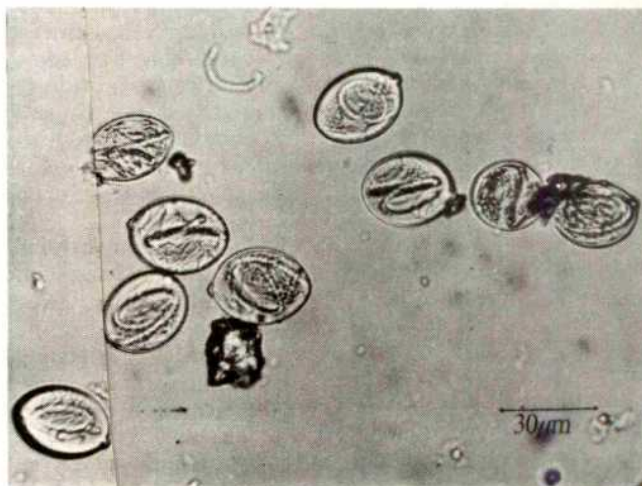

30μm

Table 1 Comparison of the effects on *A. planci* of extracts of whole coral tissue, homogenised nematocysts and mesenterial filaments

Coral	% Purity (approximate) of nematocyst preparation	Extract	Protein concentration (μg ml^{-1})	Mean response, % (to 5 drops and 10 drops, respectively)			
				Arm rearing		Tube-foot retraction	
Acropora multicaulis	60	Whole coral extract	100	8,	42	8,	33
		Homogenised nematocyst supernatant		0,	0	0,	8
Porites solida	50	Whole coral extract	73	0,	17	8,	17
		Homogenised nematocyst supernatant		0,	0	8,	8
Millepora dichotoma	85	Whole coral extract	680	0,	25	8,	25
		Homogenised nematocyst supernatant		8,	33	8,	33
Fungia fungites	—	Whole coral extract	62	25,	58	33,	67
		Mesenterial filament leachings		8,	42	8,	42
		Homogenised filament supernatant		17,	42	17,	42
Platygyra lamellina	—	Whole coral extract	85	33,	67	25,	33
		Mesenterial filament leachings		17,	42	8,	42
		Homogenised filament supernatant		8,	25	0,	42

Extracts of whole coral tissue were prepared by immersion of living corals in distilled water at 5 °C for 12 h, followed by centrifugation to remove cell debris, zooxanthellae and nematocysts, rotary evaporation to dryness, and redissolution in seawater. To isolate nematocysts, uncentrifuged whole coral extracts were fractionated by differential sedimentation to separate nematocysts from cell debris and zooxanthellae. The nematocyst preparations were homogenised in seawater to discharge them and break up the capsule, and then centrifuged. Corals were induced to extrude their mesenterial filaments by scrubbing, and the filaments were collected using forceps. They were then left overnight in filtered seawater at 5 °C, shaken to break up the structure, and centrifuged: the supernatant was tested as "leachings". The third extract was produced by homogenising the centrifugate—consisting of remaining tissue and nematocysts—followed by centrifugation. The supernatant was tested as "homogenised filament supernatant". The extracts were tested on four starfish by dripping at a rate of about one drop per s on to spots 1 cm from three arm-tips of moving starfish, avoiding dripping directly on to tube feet. The mean responses are to 5 drops and 10 drops, delivered concurrently, and were calculated by grading the strengths of responses on a four-point scale (0 to $+++$), averaging, and rounding off to the nearest whole number. As far as possible, the same animals were used for testing different extracts of the same coral species.

filament extracts are, in nearly all cases, weaker than those produced by whole coral extracts.

It thus seemed clear that neither nematocysts nor their contents were primarily responsible for the arm and tube-foot "withdrawal" responses evoked by coral extracts, although the nematocysts in living coral may have some effect. This suggested that the aversive responses constitute at least partly a response to food-related chemicals, and we conducted a search for arm-rearing activity in low-molecular-weight compounds known to be responsible for feeding attraction in other marine invertebrates[7,9]. The results are summarised in Table 2.

When the chemicals were tested individually, only betaine produced arm-rearing activity at concentrations low enough to suggest its presence in coral as an active ingredient. But a synergistic effect seems to operate when mixtures of the chemicals are tested, as observed in some other marine invertebrates[7-9].

The link between arm rearing and feeding in *A. planci* was confirmed by the finding that a mixture of betaine and three amino acids potent at inducing arm rearing— arginine, proline and glutamic acid—induced stomach

Fig. 2 Photomicrograph of a fragment of a mesenterial filament from the brain coral *Platygyra lamellina*, showing the high concentrations of nematocysts present. Two types may be distinguished where the nematocysts have been ejected from the surface of the filament by the pressure of the slide coverslip.

50μm
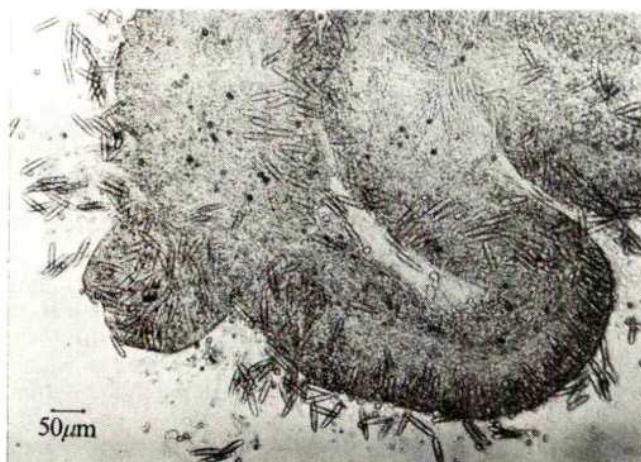
eversion in five out of nine starfish tested, using Collin's[5] assay for feeding-inducing activity.

To examine for the presence of betaine and the amino acids in coral tissue, the ninhydrin-positive fraction of the low molecular-weight fraction of an extract[4] of *Acropora* sp. was run in several sample strips alongside betaine standards and standard mixtures of the amino acids, on a paper chromatogram in a descending medium of butanol-acetic acid–water (3 : 1 : 1). When developed with a stain for alkaloids[10] it revealed a spot on the coral fraction strip running in the same position as betaine. The eluate of part of another sample strip corresponding to the alkaloid spot induced arm rearing in only two out of six trials, but the eluate of the rest of the strip possessed no activity. It thus seems likely that betaine, or another alkaloid with similar R_f in the solvent used, is the most potent factor in eliciting the arm-rearing response.

On another sample strip on the chromatogram, staining with ninhydrin revealed spots which corresponded with the position of the amino acid standards: thus their presence in coral tissue is also very likely.

Thus we found that arm rearing in *A. planci* can be induced by chemicals present in coral tissues which also stimulate it to feed. A consideration of the morphology of asteroids leads us to expect some raising of arm-tips in response to chemicals derived from prey, to aid the tube feet in location of epifaunal prey, and this is borne out by the literature[11–14]. In *A. planci*, however, the full response seems to be exaggerated beyond that of other asteroids, the arms being raised and curled back along their entire length.

Although our results suggest that nematocysts are directly responsible only to a limited extent for the aversive reaction of *A. planci* when it encounters living coral, it is certainly under threat from them, as discharged nematocysts have been found attached to tube feet after contact with coral tentacles (R.J.M., unpublished). It seems likely that the exaggerated arm curling, triggered by chemicals derived from coral tissue, serves to protect the tube feet against excessive nematocyst discharge on contact with coral. While this protection cannot be complete, the strength of the nematocyst defences of a coral species may not be as important as has been suggested (refs 1 and 2 and unpublished results of R. F. G. Ormond, H. J. Hanscomb and D. H. Beach) in determining the acceptability of that species to *A. planci* as prey.

We thank Dr S. M. Head for identification of corals, and Drs. A. C. Campbell, A. R. S. Collins and R. F. G. Ormond for helpful discussions. The work was financed by the British Ministry of Overseas Development.

RICHARD J. MOORE
CAROL J. HUXLEY

*Cambridge Coral Starfish Research Group,
PO Box 24, Port Sudan,
Democratic Republic of the Sudan*

1 Barnes, D. J., Brauer, R. W., and Jordan, M. R., *Nature*, **228**, 342–344 (1970).
2 Brauer, R. W., Jordan, M. R., and Barnes, D. J., *Nature*, **228**, 344–346 (1970).
3 Ormond, R. F. G., and Campbell, A. C., in *Proc. Second int. Coral Reef Symp.*, 1, 595–619 (Great Barrier Reef Committee, Brisbane. 1974).
4 Collins, A. R. S., *J. exp. mar. Biol. Ecol.*, **17**, 69–86 (1975).
5 Collins, A. R. S., *J. exp. mar. Biol. Ecol.*, **15**, 173–184 (1974).
6 Lowry, O. H., Rosenbrough, N. J., Farr, A. L., and Randall, R. J., *J. biol. Chem.*, **193**, 265 (1951).
7 Mackie, A. M., *Mar. Biol.*, **21**, 103–108 (1973).
8 Zafiriou, O., *Mar. Biol.*, **17**, 100–107 (1972).
9 Carr, W. E. S., *Biol. Bull. mar. biol. Lab.*, *Woods Hole*, **133**, 106–127 (1967).
10 Block, R. J., Durrum, E. L., and Sweig, G., *A Manual of Paper Chromatography and Paper Electrophoresis* (Academic, New York, 1958).
11 Valentincic, T., *Netherlands J. Sea Res.*, **7**, 191–199 (1973).
12 Fenchel, T., *Ophelia*, **2**, 223–236 (1965).
13 Christensen, A. M., *Ophelia*, **8**, 1–134 (1970).
14 Mauzey, K. P., Birkeland, C., and Dayton, P. K., *Ecology*, **49**, No. 4, 603–619 (1968).

Hybrid crown-of-thorns starfish (*Acanthaster planci* x *A. brevispinus*) reared to maturity in the laboratory

THE coral predators *Acanthaster planci* L. and *A. ellisii* (Gray) have been the subject of many publications in recent years. But there is another little known *Acanthaster* species, *A. brevispinus* Fisher, which was described from two specimens dredged in the Sulu Archipelago, Philippine Islands[1]. The species has been reported from the Great Barrier Reef region, Queensland[2,3] and specimens have been collected recently off Townsville, North Queensland, from sandy substrates and not on coral reef. Although *A. brevispinus* differs conspicuously from the other species in having very short and numerous spines on the aboral disk surface, Madsen[4] and Caso[5] were dubious of its status in their reviews of the genus *Acanthaster* Gervais. Phenotypic variation of *A. planci* was suggested in verbal discussion among Australian biologists. We describe here a resolution of the status of *A. brevispinus* by *in vitro* crosses with *A. planci* which led to the rearing of hybrid starfish.

Ripe specimens of *A. brevispinus* were collected by trawling in December 1973, near the time of annual gamete release by *A. planci* in Great Barrier Reef waters[6]. Gonads were dissected from starfish of both taxa. Ovaries were treated with 10^{-5} M 1-methyladenine in seawater to mature the oocytes[7], which were fertilised *in vitro* at 28 °C with sperm suspensions. Four groups of larvae, that is, each species and reciprocal crosses, were reared from eggs using the technique which gave a high success rate for *A. planci* larvae[8]. The different larvae passed through similar bipinnaria and brachiolaria stages. Numbers of late stage larvae in each batch represented 10–20% of the original numbers of eggs, except for the *A. brevispinus* batch in which only a few normal late brachiolariae developed. After 4 weeks, late larvae were transferred to static conditions with settlement substrates of coralline algae, coral skeleton and filamentous algae.

The newly metamorphosed starfish were five-armed and

Table 2 Arm rearing in response to food-related chemicals and mixtures

Substance	\% Arm rearing at			
	10^{-2} M	10^{-3} M	10^{-4} M	10^{-5} M
Lysine	10	0		
Arginine	40	0		
Tryptophan	70	0		
Histidine	20	0		
Hydroxyproline	0	0		
Proline	50	0		
Glutamic acid	80	0		
Betaine	—	100	65	0
Lactic acid	90	0		
Trimethylamine	100	0		
Taurine	90	0		
Trimethylamine oxide	100	0		
Mixture				
10^{-3} M Proline 10^{-3} M Arginine 10^{-3} M Glutamic acid 10^{-3} M Tryptophan 10^{-5} M Betaine		80		
10^{-4} M Proline 10^{-4} M Arginine 10^{-4} M Glutamic acid 10^{-4} M Tryptophan 10^{-6} M Betaine		30		

Solutions were tested by dripping 10 drops at about one per s about 1 cm from an arm tip of a moving starfish. Ten or 20 trials were performed for each solution and individual starfish were generally tested twice. The results are expressed as the percentage of trials in which rearing was observed.
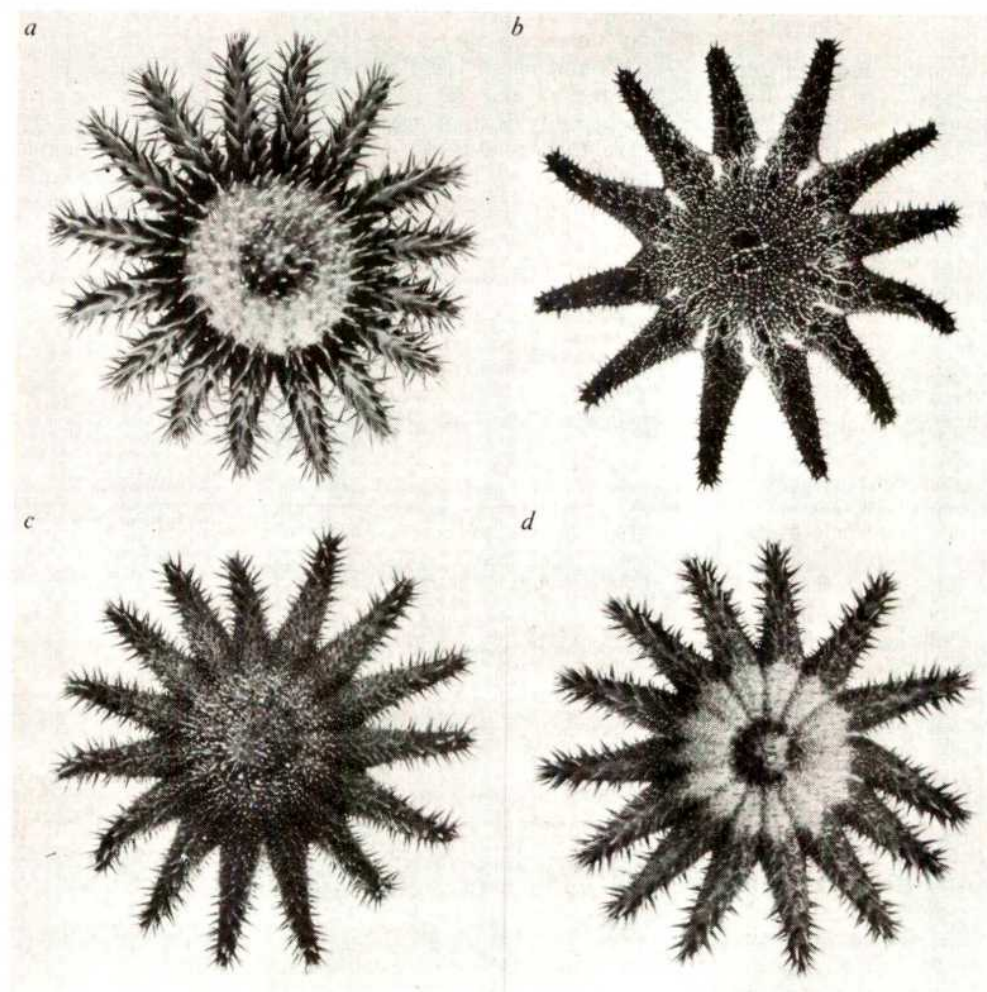
Fig. 1 Aboral views of starfish. *a*, *A. planci* (215 mm diameter); *b*, *A. brevispinus* (225 mm); *c*, *A. planci* hybrid (210 mm); *d*, *A. brevispinus* hybrid (210 mm).

0.4–1.0 mm in diameter. They fed on encrusting algae. Many failed to develop normally and, 6 weeks after metamorphosis, 60 *A. planci*, 30 *A. planci* hybrids, six *A. brevispinus* hybrids and no *A. brevispinus* remained. (For brevity, hybrids are named after the maternal taxon.) At this time, the young starfish were transferred to aquaria in a recirculating seawater system 33–35‰ salinity and 25±3 °C, and provided with coralline algae on coral skeletons for food. The three kinds of young starfish grew and added arms in the order described for *A. planci* by Yamaguchi[9]. Five to seven months after metamorphosis they achieved their adult number of arms and about this time commenced feeding on coral polyps, *Pocillopora damicornis* and *Acropora acuminata*, provided in addition to coralline algae. Late in their second year of development, the hybrid starfish and *A. planci* developed gonads with normal gametogenesis (sampled by biopsy). Sizes at 1 and 2 yr of age are shown in Table 1.

It is impossible to distinguish hybrid starfish and *A. planci* by morphology during the early months of development. But hybrids with a diameter of more than 200 mm have conspicuously shorter spines than *A. planci* and longer spines than *A. brevispinus* (Figs 1 and 2). In other features in which *A. planci* and *A. brevispinus* differ consistently,

the hybrid starfish show intermediate conditions, that is, density of aboral disk spines, length of aboral pedicellariae, length and number of adambulacral spines, shape of subambulacral spines, interradial marks on the disk and taper of arms. The intermediate hybrid features may be more similar to one parental condition, that is, on the aboral disk surface of hybrids, spine length approximates more to the *A. brevispinus* condition (Fig. 2), while the pedicellariae are slender, resembling those of *A. planci*. There are no consistent differences between *A. planci* hybrids and *A. brevispinus* hybrids. Some *A. planci* show a persistent "bull's-eye" pattern on the aboral surface from a pale ring of papulae (dermal evaginations) (Fig. 1*a*). One of three *A. brevispinus* hybrids shows the pattern (Fig. 1*d*); none of twelve *A. planci* hybrids shows the pattern.

A. planci and *A. brevispinus* differ in habitat and morphology and also in their feeding behaviour. Scallops, *Amusium balloti* and *Amusium pleuronectes*, occur in the habitat of *A. brevispinus*, and in the laboratory *A. brevispinus* traps and eats live scallops beneath the disk. It shows an arched posture when digesting live and dead scallops and scallop flesh. *A. planci* never traps live scallops and never arches when feeding on scallop flesh. (*A. planci*

Table 1 Diameters of starfish, *A. planci* and hybrids, reared in the laboratory

| | 1-yr-old | | 2-yr-old | |
	Mean diam. (mm)	Range (mm)	Mean diam. (mm)	Range (mm)
A. planci	33.4	18.0–61.2 (17)	198	145–255 (13)
A. planci hybrid	30.8	19.7–44.1 (8)	199	165–250 (9)
A. brevispinus hybrid	36.4	35.7–37.0 (2)	223	215–230 (3)

Numbers of individuals measured are given in parentheses.

Table 2 Reactions of starfish, *Acanthaster* species and hybrids, to live scallops

	Nos of Individuals	Recorded hours of activity	Mean contacts with scallop per h of activity
A. planci	3	43.6	0.50
A. planci hybrid	5	107.0	1.94*
A. brevispinus hybrid	3	62.9	2.31*
A. brevispinus	3	19.9	3.52

The three *A. brevispinus*, 220–240 mm diameter, were tested 3d after collection. The others were laboratory-reared starfish, 160–190 mm diameter, which had been fed solely with algae and coral before this experiment. Each starfish was placed in a glass aquarium 30 × 30 × 50 cm with running seawater and continuous illumination. Activity was recorded by an 8-mm movie camera, situated above the aquarium, making exposures at 30-s intervals. The starfish was accustomed to the apparatus for a day and then a live scallop, *Amusium pleuronectes*, was added and filming commenced. Duration of filming was 27–31 h. Film strips were analysed frame by frame. Each new contact with the scallop by the starfish, that is, passing at least part of an arm over the scallop, and periods of quiescence were recorded. Data for each starfish were recorded as numbers of contacts per unit period (h) of locomotion in case contacts with scallops were purely incidental, being related only to the amount of activity. Numbers of 0, 1 . . . 6+ contacts per hour were accumulated for each kind of starfish. For statistical analyses, these numbers were adjusted to 100 h total activity and the numbers in each category for two kinds of starfish compared by χ^2 analysis.
*These data are not significantly different ($P < 0.5$); they are the only data for two kinds of starfish which are not significantly different ($P < 0.001$) (see explanation of statistical analysis above).

normally feeds on hermatypic coral, but it will accept a range of animal flesh when hungry.) The hybrid starfish trap scallops and they may adopt an arched posture, although not as pronounced as *A. brevispinus,* when feeding on scallops. Further evidence of behavioural reactions to scallops by hybrid starfish, inherited from their *A. brevispinus* parent, is shown in Table 2. Hybrid starfish reared on coral showed a significantly greater rate of contacts with scallops than *A. planci* of similar experience. Contacts with the scallop usually occurred when the starfish moved towards it across the floor of the aquarium and contact was often terminated by the scallop escaping with rapid valve movements.

It has been shown that the starfish described as *A. brevispinus* are genetically different from typical *A. planci* : crossing them produces starfish which differ in morphology and behaviour from *A. planci* reared in the same conditions. There is no evidence of gamete incompatibility between the two taxa, as found in interspecific crosses attempted with other asteroids[10,11]. There is obviously a high degree of genetic compatibility between them, at least to F_1 hybrids. But they are isolated in the field by their habitat preferences, *A. planci* on coral reefs and *A. brevispinus* in deeper water on sandy substrates. Although probably they both spawn during summer in Great Barrier Reef waters, the chance of cross fertilisation is remote: ripe *A. planci* spawn in response to a pheromone released with gametes by neighbouring individuals[12] and this mechanism is unlikely to operate over hundreds or thousands of metres, with consequent dilution of the pheromone and gametes. No natural intermediate forms have been found in the Great Barrier Reef region,

Fig. 2 Maximum disk (*a*) and arm (*b*) spine lengths on the aboral surfaces of *Acanthaster planci* (●), *A. brevispinus* (▲), *A. planci* hybrids (○) and *A. brevispinus* hybrids (△). Regression lines were fitted by least squares method (hybrid data pooled).

although the reef environment has been surveyed extensively by biologists aware of the morphology of *A. brevispinus*[13,14].
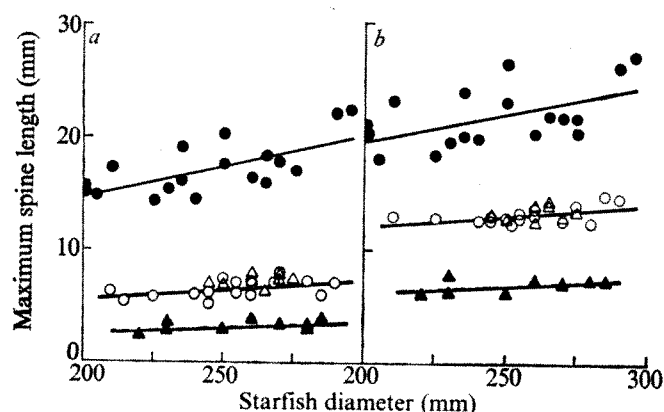
A. planci and *A. brevispinus* are taxa which do not interbreed in a region of sympatry. They are species retaining a high degree of genetic compatibility, suggesting recent speciation. It is notable that the morphological differences between *A. planci* and *A. brevispinus* are much greater than the morphological variation of *A. planci* throughout its wide Indo-Pacific distribution and between *A. planci* and the eastern Pacific taxon, *A. ellisii,* which is of disputed status[15,16]. If degree of morphological difference within this genus even approximately reflects degree of genetic difference, then all populations of *A. planci* and *A. ellisii* are genetically compatible. These suggestions are in accordance with a recent evolution of *A. planci* and dispersal through the Indo-Pacific region. It seems more likely that *A. planci* evolved from an ancestor like *A. brevispinus* than vice versa. *A. planci* is atypical of coral reef asteroids in its large size and often conspicuous mode of life. It is more probable that an atypical, specialised, coral predator evolved from an unspecialised, sand inhabitant than vice versa. In this case, development of large sharp spines from the small blunt spines of *A. brevispinus* gave the means of protection from the numerous predators which are a particular hazard of the habitat and mode of life adopted by *A. planci.*

This is the first time that asteroid hybrids have been reared successfully in the laboratory. The only other reported cross fertilisations resulted in abnormal embryos or malformed early larvae[10]. In contrast, there is an extensive literature on echinoid hybridisation with crosses involving quite divergent taxa[17]. Although almost no echinoid hybrids have been reared beyond plutei, hybrids from reciprocal crosses between the sand dollars *Dendraster excentricus* and *Encope californicus* have been reared to young urchins and in these hybrids paternal characters predominate[17]. In *A. planci* and *A. brevispinus* hybrids, characters of neither parent predominate and the hybrids generally show intermediate conditions.

J. S. Lucas
M. M. Jones

*Zoology Department,
James Cook University of North Queensland,
Queensland, 4811, Australia*

Received June 9; accepted August 23, 1976.

1 Fisher, W. K., *Proc. biol. Soc. Wash.*, 30, 89–93 (1917); *Bull. U.S. natn. Mus.*, 100, xi + 711 (1919).
2 Bennett, I., *Proc. Linn. Soc. N.S.W.*, 83, 375–376 (1958).
3 Endean, R., *Pap. Dept Zool. Univ. Qd*, 1, 289–298 (1961).
4 Madsen, F. J., *Vidensk. Meddr Dansk naturh. Foren.*, 117, 179–192 (1955).
5 Caso, M. E., *Ann. Inst. Biol. Univ. Mex.*, 32, 313–331 (1961).
6 Lucas, J. S., *Micronesica*, 9, 197–203 (1973).
7 Kanatani, H., *Expl Cell Res.*, 57, 333–337 (1969).
8 Lucas, J. S., in *Crown-of-thorns Starfish Seminar Proc.*, Brisbane, September 1974, 109–121 (Australian Government Publishing Service, Canberra, 1975).
9 Yamaguchi, M., in *Biology and Geology of Coral Reefs*, 2, *Biology*, 1 (edit. by Jones, O. A., and Endean, R.), 369–387 (Academic, New York, 1973).
10 Gemmill, J. F., *Trans. zool. Soc., Lond.*, 20, 1–71 (1912).
11 Tyler, A., and Tyler, B. S., in *Physiology of Echinodermata* (edit. by Boolootian, R. A.), 683–741 (Interscience, New York, 1966).
12 Beach, D. H., Hanscomb, N. J., and Ormond, R. F. G., *Nature*, 254, 135–136 (1975).
13 Pearson, R. G., *Nature*, 237, 175–176 (1972).
14 Endean, R., and Stablum, W., *Environ. Conserv.*, 2, 247–256 (1975).
15 Glynn, P. W., *Science*, 180, 504–506 (1973).
16 Barham, E. G., Gowdy, R. W., and Wolfson, F. H., *Fish. Bull.*, 71, 927–942 (1973).
17 Hinegardner, R. T., *Am. Zool.*, 15, 679–689 (1975).

Cultivation of larvae of Japanese eel

ALTHOUGH the breeding place of the European eel has been detected[1], the morphology and ecology of their larvae, especially in pre-leptocephalus stages, still remain to be investigated. We have, however, succeeded[2-4] in obtaining some pre-larvae of the Japanese eel, *Anguilla japonica*, following artificial induction of maturation and spawning in aquaria. The larvae survived for only 6 d; their teeth characteristic of muraenoid larvae were merely in the state of anlagen and their eyes were completely devoid of retinal pigment. Further information of the early development of the eel is therefore required to enable elucidation of their life history. Here, we describe the larvae of Japanese eel developing for up to 14 d in the laboratory.

Silver eels were collected in the Hiranuma and Mabuchi Rivers in Aomori prefecture (Japan) in September 1975. Females and males were matured with repeated injections of chum salmon pituitary homogenate and Synahorin (Teikoku Zoki), respectively[5,6]. Eggs were stripped from the females, placed into glass dishes and inseminated by the dry method with the sperm of 2 or 3 males. The eggs were kept in seawater at 23 °C until they hatched. The larvae were reared at 23 °C on the day of hatching. Subsequently the larvae were maintained at 19 °C.

The larvae survived for 14 d. The following description is based on the observation of the larvae from days 7 to 14 after hatching. Since larvae of that period of development are very weak and their tail is liable to shrink a little during microscopic observations, the measured values of body length and the myomere number of larvae

Fig. 1*a*, Head of larva, day 9 after hatching, × 25; *b*, head of larva, day 12 after hatching, × 22; *c*, larva, day 14 after hatching, × 13.5; *d*, head of larva, day 14 after hatching, × 38.5.

may be approximate. On day 7, the mean body length was ~6.2 mm, and on day 14, ~7.0 mm. The number of myomeres was 53–54 + 50 (pre- + postanal myomeres) 7–10 days after hatching and 54 + 55–60 in those 11–14 days after hatching.

The larvae on day 7 had well developed tooth anlagen on the jaws and a large ventriculus cerebri in the brain. A broad embryonic fin enveloped the compressed body continuously from the posterior region of the head to the caudal margin of the vent. There was no trace of pigmentation except in the caudalmost part of the membraneous fin. By day 8, some black retinal pigment began to appear in both eyes (Fig. 1*a*) and after a further 12 h pigmentation of the eyes was fully developed. During that time the ventriculus cerebri became smaller. These larvae, unlike the earlier stages[4], were seen to swim in a horizontal position and to rest suspended in the water in a head-down attitude. From days 9 to 11 after hatching, both the upper and lower jaws developed considerably, and the tooth anlagen came to differentiate into definite teeth, which lengthened gradually, one after another. By day 12, sharp teeth were observed to protrude obliquely from the jaws (Fig. 1*b*). An oil drop near the cranial end of the yolk sac became extremely small. On day 14 after hatching, black pigmentation was still seen in the larvae, but only in the eyes and on the caudalmost tip of the membraneous fin (Fig. 1*c*). A pair of pectoral fins were present. The mouth which was directed downwards during the foregoing stages took a forward direction. The upper jaw had three pairs of shorter teeth in addition to the longest, grasping tooth, while the lower jaw had four pairs of teeth (Fig. 1*d*). The lower jaw moved occasionally, but the mouth did not close perfectly.

The most advanced larvae obtained seem to resemble closely the smallest European eel larvae collected by Schmidt[7] in the Atlantic Ocean. From these findings on both the European and the Japanese eel, a clearer outline of the morphological changes of the eel during the initial stage of its life cycle has been obtained.

K. YAMAUCHI
M. NAKAMURA
H. TAKAHASHI
K. TAKANO

Department of Biology,
Faculty of Fisheries,
Hokkaido University,
Hakodate, 040, Japan

1 Schmidt, J., *Phil. Trans.*, B211, 179–208 (1922).
2 Yamamoto, K., and Yamauchi, K., *Nature*, 251, 220–222 (1974).
3 Yamamoto, K., Yamauchi, K., and Kasuga, S., *Bull. Jap. Soc. scient. Fish.*, 41, 21–28 (1975).
4 Yamamoto, K., Yamauchi, K., and Morioka, T., *Bull. Jap. Soc. scient. Fish.*, 41, 29–34 (1975).
5 Yamamoto, K., Hiroi, O., Hirano, T., and Morioka, T., *Bull. Jap. Soc. scient. Fish.*, 38, 1083–1090 (1972).
6 Yamamoto, K., Morioka, T., Hiroi, O., and Omori, M., *Bull. Jap. Soc. scient. Fish.*, 40, 1–7 (1974).
7 Schmidt, J., *Carlsberg Foundation, Copenhagen* (1935).
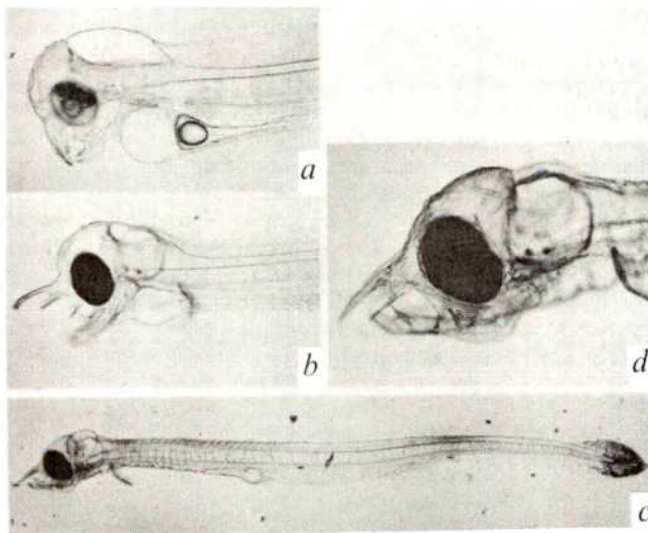
Genetics of expression of xenotropic virus and autoimmunity in NZB mice

CERTAIN retroviruses (RNA tumour viruses) have been implicated in the autoimmune disease of New Zealand mice[1]. These mice produce large numbers of xenotropic retroviruses[2,3] and contain high concentrations of the retroviral envelope antigen gp 69/71 in their serum and tissues[4]. Moreover, gp 69/71 and the corresponding antibodies contribute to the immune deposits in the nephritic kidneys of NZB and (NZB × NZW)F₁ mice[4]. Nevertheless, it is not established that xenotropic viruses are the primary cause of the autoimmune disease of NZB mice. Conceivably, these agents may be involved only secondarily. The presence of RNA viruses in NZB mice may explain neither the production of antibodies against DNA, nucleoproteins, and erythrocyte antigens nor the anomalies of T- and B-lymphocyte function[5]. Transmission of autoimmunity has not been achieved
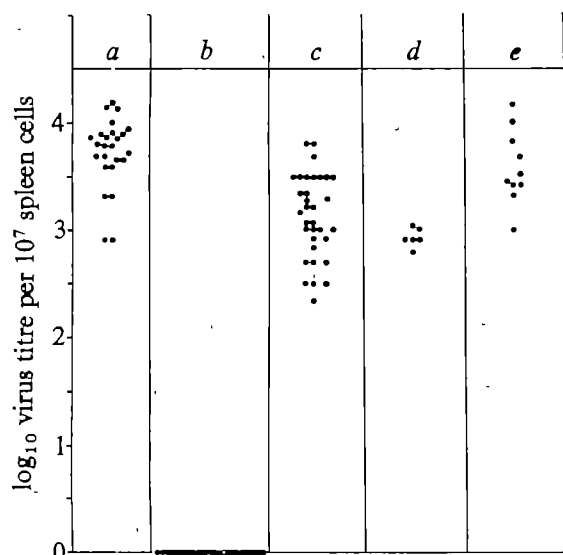
Fig. 1 Xenotropic virus titres in NZB and SWR mice and their crosses. All parental strains of mice were obtained from Jackson Laboratory, Bar Harbor, Maine. Spleen cell suspensions were prepared from the test animals as described previously[10] and assayed on the 81 cell line[11] for xenotropic virus as follows. The 81 cells were grown and the assay was carried out in McCoy's 5A medium (Grand Island), supplemented with antibiotics and 15% foetal calf serum. The 81 cells (2×10⁶) were seeded into 6-cm plastic Petri dishes. The next day the medium was withdrawn and the cells were treated with 2 ml DEAE dextran in McCoy's 5A medium without serum (25 μg ml⁻¹) for 30 min at 37 °C. Serial dilutions of spleen cell suspensions of 5 ml McCoy's 5A medium were inoculated on the 81 cells after the DEAE dextran was rinsed away. On days 3 and 10 after inoculation 3 ml medium was added (similar results were obtained when on day 0 the spleen cells were inoculated in 10 ml medium and the cultures left undisturbed until day 7). On day 7 the medium was changed completely. Foci were counted on unstained dishes on days 13–14 after infection. Titres are expressed in this and succeeding Figs as log₁₀ of the number of foci induced per 10⁷ inoculated spleen cells. In experiments not shown here we found that the number of foci that developed increased linearly and proportionally to the number of virus-producing spleen cells inoculated. Satellite foci did not occur in the dishes containing lower dilutions of the virus-producing spleen cells. Each circle represents an individual animal. ●, Male; ○, female. *a*, NZB (1–9 months); *b*, SWR (2–11 months); *c*, (SWR×NZB)F₁ (1–9 months); *d*, (NZB×SWR)F₁ (2 weeks); *e*, (SWR×NZB) F₁×NZB (3–5 months).

with cell-free filtrates from NZB mice[6], nor by transplanting NZB lymphomas[7] which are known to produce large numbers of virus particles[8]. Only spleen cells from old NZB mice can transmit the autoimmune disorder[6]. Transmission of the disease with cell-free extracts is unlikely to succeed because the NZB xenotropic virus cannot productively infect mouse cells; it can infect only cells of heterologous species[1,2]. Here we deal

with the vertical transmission of NZB viral genes as Mendelian traits—a phenomenon that has been demonstrated with other retroviruses[9]—to test the hypothesis that the development of autoimmunity in NZB mice and their crosses is independent of the expression of high titres of xenotropic viruses. We have found that the high grade expression of infectious xenotropic virus characteristic of NZB mice is a genetically determined trait. Two independently segregating loci seem to specify the NZB phenotype. So far, animals that are virologically like NZB have failed to develop signs of autoimmune disease.

Infectious xenotropic virus was measured by a modification[10] of the test described by Fischinger *et al.*[11]. Spleen cells from the test animal were cocultivated on monolayers of the 81 clone of the murine sarcoma virus-transformed CCC feline cell line (provided by Dr Peter Fischinger)[11]. Xenotropic viruses of various origins replicate in the 81 cell line and produce foci which are proportional in number to the input of virus-producing cells. This quantitative assay detected xenotropic viruses produced by cells of different mice, including BALB/c, B10.A, C57BL/6, AKR and C57L (Fig. 2 and unpublished observations). The assay reflected the virological status of the spleen cells on the day they were inoculated on the 81 cell line because mitomycin treatment of NZB spleen cells just before inoculation did not reduce the titre of xenotropic virus relative to that of untreated NZB spleen cells. Thus, the assay reflected the *in vivo* virological status (phenotype) of the animal, as was demonstrated in the case of ecotropic virus-producing strains[10].

Figure 1 shows the expression of xenotropic virus in two strains, NZB and SWR. As found by others[3,4], xenotropic virus was present in high titres in all the NZB mice (age range 1–9 months), whereas in none of the 25 SWR mice (2–11 months) could xenotropic virus be detected. It seems that the production of infectious xenotropic virus by SWR mice is either 'switched off' or proceeds at an extremely low rate. Alternatively, SWR mice may produce a defective virus. In any case, the 81 cell line assay clearly distinguished between NZB and SWR phenotypes: the former were uniformly positive in high titre and the latter uniformly negative.

Figure 1 also shows that all (SWR×NZB)F₁, (NZB×SWR)F₁ and (SWR×NZB)F₁×NZB backcross mice expressed high titres of xenotropic virus early in life. This indicates inheritance of the NZB virus phenotype by a genetic mechanism that is dominant, highly penetrant and contributed equally well by both sexes. No linkage to the sex of the animal was detected. We also tested the expression of infectious xenotropic virus in other crosses: (C57BL/6×NZB)F₁, (B10.A×NZB)F₁ and (AKR×NZB)F₁ (Figure 2). In all cases a picture similar to the (SWR×NZB)F₁ cross emerged—dominant, highly penetrant expression of the NZB viral phenotype.

Tests of the first backcross with SWR and the F₂ generations showed clearcut segregation that was independent of the maternal direction of the cross (Fig. 3). In the SWR backcross progeny, the ratio of virus-positive to virus-negative mice was

Table 1 Autoimmune markers and disease in NZB mice and their crosses with C57BL/6(B6), SWR and B10.A mice

Strain	Age (months)	Direct Coombs' test	Anti-DNA antibody		Nephritis
			Native	Denatured	
(B6×NZB)F₁ male	6–7		0/6	0/6	0/14
	9	0/9	0/14	0/14	0/14
(SWR×NZB)F₁ male	6–7		0/14	4/14*	0/18
	8	0/8	0/8	1/8*	0/8
(B10.A×NZB)F₁ male	8	0/4	0/4	0/4	0/13
female	7		0/7	1/7*	0/7
NZB male	8–12	6/8	3/8	8/8**	3/8
	3–5	0/7	0/7	7/7**	0/7

Direct Coombs' tests were carried out according to the method of Long *et al.*[15]. Antibodies to native and denatured DNA in the sera of the mice was quantified by the ¹⁴C DNA/glass filter technique developed by Lewis *et al.*[16]. Percentage binding of denatured DNA in the positive serum samples of the F₁ hybrids* was in the range 10–18% (mean 12%), whereas among the NZB mice** it was 14–36% (mean 25%).
Nephritis: degree of glomerular involvement was determined by histological criteria of Tonietti *et al.*[17], using haematoxylin–eosin and PAS-stained sections of kidneys.

approximately 3:1 (51:18, $P > 0.75$). This suggests two independently segregating loci either of which gave a positive phenotype. According to this interpretation, the virus-negative mice were the result of simultaneous homozygosity of recessive alleles to these two autosomal genes. Results obtained in the F_2 generation tended to support this interpretation, as a 15:1 ratio (54:4, $P > 0.9$) was obtained. So far, no linkage between coat colour and virus expression has been found. Linkage with other traits is being sought.

The quantitative data further support a multiple independent gene model because intermediate titres of virus (up to 150 focus-forming units (FFU)) were found in some mice of the backcross and F_2 generations. In contrast, all the F_1 mice expressed high titres of virus. The segregation ratios for high:intermediate:negative virus phenotype in the backcross and F_2 mice suggest that one of the NZB loci is responsible for high grade virus expression when present alone or in association with the other virus-inducing locus. Let us call the first locus NZV_1 and the second NZV_2 for convenience. The ratios suggest that when NZV_2 is present by itself, intermediate expression of virus would result. This hypothesis is being tested by crossing intermediate-titred mice with each other or with SWR mice to determine if the effect is caused by a segregating locus. If this is so, the hypothesis of phenotypic mixing between the NZB virus and a presumed defective SWR virus to give intermediate titres can be rejected. It is also conceivable that the high and intermediate titres are the result of two different types of NZB xenotropic viruses. This is being analysed.

The virological patterns in the various F_1 hybrids were quantitatively and qualitatively similar, indicating that no major genetic influences inhibited the expression of NZB virus. There was a minor effect in that $(B10.A \times NZB)F_1$ hybrids tended in general to have lower titres than the other two hybrids.

Although the data fit a two-locus model and are incompatible with a single-locus model, we cannot exclude that more than two virus-inducing loci are segregating. If there were one locus with high penetrance and several other loci with low penetrance, the overall segregation ratios might mimic a two-locus pattern. Further studies of backcross families will clarify this point.

We cannot state whether the postulated NZB loci are genetic elements of the virus itself or are genes that promote the expression of the virus. The distinction between these alternatives requires the use of genetic markers that distinguish
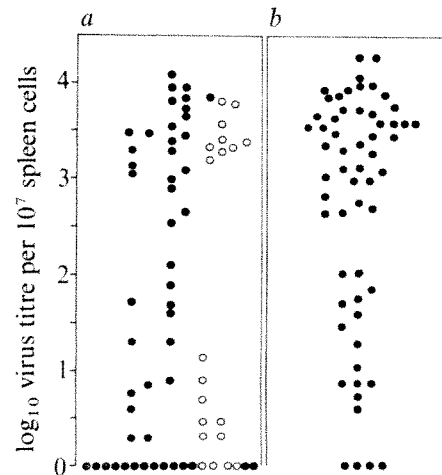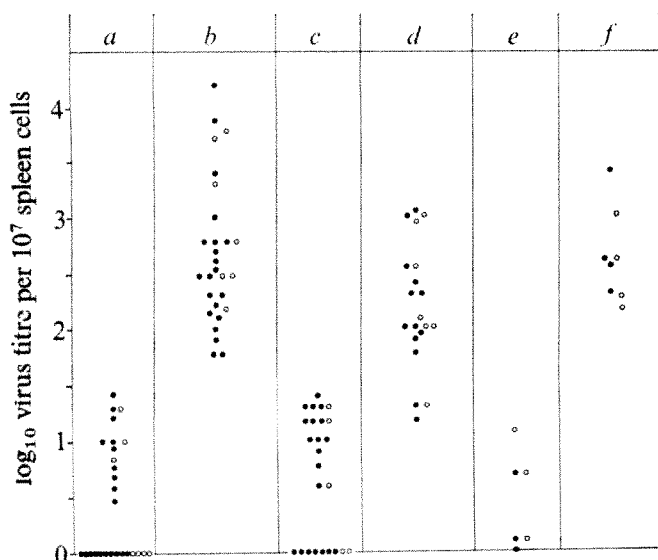

Fig. 3 Xenotropic virus titres in mice of first backcross with SWR and the F_2 generations. All mice were 3–4 months old. a: ●, SWR × (SWR × NZB)F_1; ○, (SWR × NZB)F_1 × SWR. b, (SWR × NZB)F_2.

between subtypes of xenotropic viruses (for example, NZB-xenotropic virus, or MuLV-x$^\beta$, and C57BL/6-xenotropic virus, or MuLV-x$^\alpha$)[12]. It is likely that we are not dealing with either the presence or absence of genetic information required for the phenotype under study since the low virus strains C57BL/6, B10.A and AKR obviously can make xenotropic virus; and the SWR mouse, although 'virus negative' when tested in the 81 (cat) and A673 (human)[13] cell lines, is known to have the p12 of xenotropic virions in its cells[13].

Table 1 shows the results of preliminary studies for autoimmune disease in NZB mice and their crosses with SWR, C57BL/6 and B10.A mice. Markers of autoimmunity were readily detected in NZB mice. In contrast, these abnormalities were either absent or infrequent in all of the F_1 crosses tested, although all the animals expressed high titres of xenotropic viruses from an early age. These data do not take into account the possibility that autoimmune markers or autoimmune disease may develop later in the F_1 hybrids than in the NZB parent. This time-dependent phenomenon has been noted by others and is a complicating feature of genetic analysis of autoimmunisation in NZB mice and their crosses[14]. Since the material required for virological analysis was obtained in all of the mice by biopsy of the spleen, these animals are currently being aged and will be retested for signs of autoimmunisation at a later time. It will also be of interest to note if the virus-negative backcross and F_2 progeny develop autoimmune disease later in life. Such a finding would support the hypothesis that the autoimmune abnormalities of NZB mice are independent of xenotropic viruses.

SYAMAL K. DATTA
ROBERT S. SCHWARTZ

Hematology Service,
New England Medical Center Hospital, and
Department of Medicine,
Tufts University School of Medicine,
Boston, Massachusetts 02111

1 Mellors, R. C., Shirai, T., Aoki, T., Huebner, R. J., and Krawczynski, K., *J. exp. Med.*, **133**, 113–132 (1971).
2 Levy, J. A., *Science*, **182**, 1151–1153 (1973).
3 Levy, J. A., Kazan, P., Varnier, O., and Kleiman, H., *J. Virol.*, **16**, 844–853 (1975).
4 Yoshiki, T., Mellors, R. C., Strand, M., and August, J. T., *J. exp. Med.*, **140**, 1011–1027 (1974).
5 Talal, N., *Progr. clin. Immun.*, **2**, 101–120 (1974).
6 Burnet, F. M., *Progress in Immunology*, **5** (edit. by Brent, L., and Holborow, J.), 27–36 (North-Holland, Amsterdam, 1974).
7 East, J., *Prog. exp. Tumor Res.*, **13**, 84–134 (1970).
8 Mellors, R. C., and Huang, C. Y., *J. exp. Med.*, **124**, 1031–1038 (1966).
9 Rowe, W. P., *J. exp. Med.*, **136**, 1272–1285 (1972).

Fig. 2 Xenotropic virus titres in C57BL/6, B10.A and AKR mice and their F_1 hybrids bred by crossing with NZB. ●, Male; ○, female. a, C57BL/6 (8–9 months); b, (C57BL/6 × NZB)F_1 (6–9 months); c, B10.A (5–9 months); d, (B10.A × NZB)F_1 (5–8 months); e, AKR (4–6 months); f, (AKR × NZB)F_1 (2–7 months).

10 Melief, C. J. M., *et al.*, *Proc. Soc. exp. Biol. Med.*, **149**, 1015–1018 (1975).
11 Fischinger, P. J., Blevins, C. S., and Nomura, S., *J. Virol.*, **14**, 177–179 (1974).
12 Callahan, R., Lieber, M. M., and Todaro, G. J., *J. Virol.*, **15**, 1378–1384 (1975).
13 Stephenson, J. R., Reynolds, R. K., Tronick, S. R., and Aaronson, S. A., *Virology*, **67**, 404–414 (1975).
14 Ghaffar, A., and Playfair, J. H. L., *Clin. exp. Immun.*, **8**, 479–490 (1971).
15 Long, G., Holmes, M. C., and Burnet, F. M., *Aust. J. exp. Biol. Med. Sci.*, **41**, 315–322 (1963).
16 Lewis, R. M., Stollar, B. D., and Goldberg, E. B., *J. Immun. Meth.*, **3**, 365–374 (1973).
17 Tonietti, G., Oldstone, M. B. A., and Dixon, F. J., *J. exp. Med.*, **132**, 89–109 (1970).

Sex differences in formation of anti-T-cell antibodies

A NATURALLY occurring thymocytotoxic autoantibody (NTA) has been found in significant titre early in life in NZB mice[1]. This antibody is present in almost all NZB mice by 3 months of age, and is associated with an age-dependent loss of T cells in these mice[2-4]. NTA may be responsible for an initial loss of suppressor T cells in NZB mice and thereby contribute to the development of autoimmunity. Although other mouse strains (NZW, C57BL/6J, AKR/J, BALB/cJ, and 129) may produce NTA later in life, the prevalence and titre of NTA is much lower than in NZB mice. NTA reacts with brain- and thymus-derived cells as well as thymocytes, but not B cells or non-lymphoid abdominal organs[1-5]. In contrast to standard anti-θ serum, NTA can combine with both Thy 1.1 (θ-AKR) and Thy 1.2 (θ-C3H)[1,6]. Genetic factors have been implicated in the development of autoimmunity in New Zealand mice[7,8]. In addition, X-linked immune response genes have been described for synthetic[9] and naturally occurring[10] nucleic acids. Moreover, female hybrids (NZB/NZW) develop a more severe disease than males. We therefore undertook the present study to determine the contribution of the X chromosome to the development of NTA.

Hybrid mice were produced by crossing NZB/N and DBA/2 strains. The DBA/2 strain was chosen because of the reported lack of warm-reacting NTA at approximately 1 yr of age[2]. Since the presence of NTA in high titres occurs earlier in the life of NZB than in other strains, 12-month-old hybrid mice were chosen for study so that differences in the formation of NTA would not represent time differences alone, but genetic background as well. Because the X chromosome of the NZB strain may be important in determining the presence of NTA, we wished to rule out a simple hormonal mechanism. To study this (DBA/2 × NZB/N) and (NZB/N × DBA/2)F₁ hybrid virgin mice of both sexes were castrated at 5 weeks of age[11], with littermates left intact. Sera obtained at 1 yr of age were studied for the presence of NTA as described previously, with only minor modifications[6]. Briefly, thymocytes from 4–6-week-old C57BL/6N were incubated with 30 μCi ⁵¹Cr per 10⁷ cells for 30 min at 37 °C. Labelled thymocytes (5 × 10⁴) were added to 50 μl of serially diluted test serum and incubated for 30 min at room temperature followed by 30 min at 4 °C. Cells were

spun, washed twice and incubated with 50 μl 1:15 rabbit complement (previously absorbed with C57BL/6N thymocytes) for 30 min at 37 °C. The suspension was then centrifuged and 50 μl of supernatant removed and counted. The positive control was a pool of NZB sera with known NTA activity; the negative control consisted of sera from 1-yr-old female DBA/2 mice. Maximum ⁵¹Cr release was obtained by freeze-thawing, and cytotoxicity of complement (C) alone determined. Percentage cytotoxicity was calculated as

$$\frac{(\text{c.p.m. test serum} - \text{c.p.m. C})}{(\text{c.p.m. freeze-thaw} - \text{c.p.m. C})} \times 100$$

All sera were assayed in duplicate with good reproducibility. The cytotoxic titre was the last dilution giving greater than 50% ⁵¹Cr release. A positive titre was at least 1:4.

Approximately half of the unmanipulated female NZB × DBA mice had NTA at 1 yr of age, whereas none of the males was positive (Table 1). Castration of the females did not affect the percentage positive or the titre of NTA. Castration of the males at 5 weeks of age, however, led to NTA formation indistinguishable from that of females (Table 1).

DBA-mothered F₁ hybrids had a lower incidence of NTA than corresponding NZB-mothered mice. Nevertheless, the same pattern of NTA incidence was observed. Hybrid males failed to produce NTA unless they were castrated.

The studies of uncastrated mice demonstrate NTA formation in female but not male F₁ hybrids. Superficially, the results for the DBA-mothered F₁ hybrids might suggest a role for an immune response gene carried on the NZB X chromosome, but not the DBA X chromosome. This would be in agreement with a previous study of the anti-DNA response[10]. According to this model, NZB-mothered F₁ males would receive a 'responder' X chromosome and produce NTA. We found, however, that NZB-mothered F₁ males failed to produce NTA, indicating a mechanism other than an X-linked immune response gene. Moreover, the observation that castration abolished the difference between hybrid males and females, suggested a sex hormonal mechanism. Thus, the lack of NTA in male NZB hybrids is due in part to a functional testis; the possibility that male sex hormones are responsible is under investigation. It is not clear why NZB males develop NTA. They are relatively infertile and may have a deficiency in suppressive male sex hormones. Alternatively, additional factors (such as a more rapid loss of suppressor cells than occurs in NZB/DBA male F₁ hybrids) may operate to make it possible for them to develop NTA. In addition, both NZB male and female mice share similar severity of disease manifestations unlike the NZB/NZW F₁ model where females clinically have a more severe expression of autoimmunity[8,11].

The data reported here provide a possible explanation for the female predominance in both murine and human autoimmunity. Male sex hormones may inhibit the production of certain autoantibodies; therefore, males would be protected from the development of autoimmunity. The study of such a mechanism may lead to a better understanding of autoimmunity as well as potential therapeutic intervention.

ELIZABETH S. RAVECHE

Laboratory of Experimental Pathology,
Section on Cytogenetics,

LYNELL W. KLASSEN
ALFRED D. STEINBERG

Arthritis and Rheumatism Branch,
National Institute of Arthritis,
Metabolism and Digestive Diseases,
National Institutes of Health,
Bethesda, Maryland 20014

1 Shirai, T., and Mellors, R. C., *Proc. natn. Acad. Sci. U.S.A.*, **68**, 1412–1415 (1971).
2 Shirai, T., and Mellors, R. C., *Clin. exp. Immun.*, **12**, 133–152 (1972).

Table 1 NTA at 1 yr of age

F₁ hybrid	Sex	Positive NTA test Number/total	%
♀ ♂			
NZB × DBA	♀	11/20	55
NZB × DBA	♀, castrated	9/16	56
NZB × DBA	♂	0/15	0
NZB × DBA	♂, castrated	10/18	56
DBA × NZB	♀	2/8	25
DBA × NZB	♀, castrated	2/7	29
DBA × NZB	♂	0/6	0
DBA × NZB	♂, castrated	2/7	29
Parental strain			
NZB	♀	10/10	100
NZB	♂	10/10	100
DBA	♀	0/8	0
DBA	♂	0/8	0

[3] Shirai, T., Yoshiki, T., and Mellors, R. C., *Clin. exp. Immun.*, **12**, 455–464 (1972).
[4] Shirai, T., Yoshiki, T., and Mellors, R. C., *J. Immun.*, **110**, 517–523 (1973).
[5] Golfand, M. C., Parker, L. M., and Steinberg, A. D., *J. Immun.*, **113**, 1–8 (1974).
[6] Parker, L. M., Chused, T. M., and Steinberg, A. D., *J. Immun.*, **112**, 285–292 (1974).
[7] Howie, J. B., and Helyer, B. J., *Ann. N. Y. Acad. Sci.*, **124**, 167–177 (1965).
[8] Talal, N., and Steinberg, A. D., *Current Topics in Microbiology and Immunology*, **64**, 79–103 (Springer, New York, 1974).
[9] Scher, I., Frantz, M. M., and Steinberg, A. D., *J. Immun.*, **110**, 1396–1401 (1973).
[10] Mozes, E., and Fuchs, S., *Nature*, **249**, 167–168 (1974).
[11] Steinberg, A. D., Pincus, T., and Talal, N., *Immunology*, **20**, 523–531 (1971).

Serum containing endotoxin-induced tumour necrosis factor substitutes for helper T cells

WHEREAS immunocompetent precursor B cells recognise antigenic determinants by way of specific surface receptors (an interaction which presumably initiates cell division[1,2]), the binding of antigen to surface receptors alone does not result in the differentiation of B cells to antibody-secreting cells. To accomplish this differentiation step a second signal is needed. This signal may be provided by T cells in the form of a soluble factor which they produce *in vitro* after contact with antigen[3-5]. Besides T cells, activated peritoneal macrophages have also been shown to produce a factor which allows B cells from athymic Nu/Nu mice to produce antibody against fowl gammaglobulin *in vitro*. Production of this factor is increased by addition of bacterial lipopolysaccharide (LPS) to the culture medium[6]. LPS itself has been shown to provide the second signal (antigen being the first) for antibody production by B cells[7], but it is not clear whether the effect on antigen-reactive B cells is direct, or mediated by other cells. We report here that injection of LPS, in certain conditions in the mouse, induces the release of factors into the serum which enable B cells to produce antibody to heterologous red-cell antigen *in vitro* in the absence of helper T cells. The factor is released in large quantity when LPS is injected into mice that have been primed with BCG. It is hardly, if at all, detectable in the serum of mice injected with LPS but not pretreated with BCG. Our investigation was prompted by the discovery that serum from BCG-infected mice given LPS causes acute necrosis of established tumour grafts in syngeneic mice (tumour necrosis serum, TNS)[8]. After partial purification, tumour-necrotising activity was shown to reside in a glycoprotein fraction (termed tumour necrosis factor, TNF) with a molecular weight of about 150,000 that migrates with the α-2-globulins[9].

TNS used in the experiments reported here was produced in CD-1 Swiss mice. Twenty million living BCG organisms (Trudeau Institute, Saranac Lake, New York) were injected intravenously, followed by intravenous injection of 25 μg LPS (Difco) at the peak of hepatosplenomegaly, usually 2 weeks later. The mice were exsanguinated 2 h after injection and their serum separated and passed through a Millipore filter before use in tissue culture.

The production of haemolytic antibody by spleen cells from BDF[1] mice was studied in the Mishell–Dutton culture system[10]. Initial experiments showed that the production of antibody to sheep red blood cells (SRBCs) by unfractionated spleen cells was accelerated and augmented by addition of TNS or TNF to the culture medium (Table 1). This led to the question of whether TNF acted directly on B cells, or indirectly through macrophages or T cells. Experiments with fractionated spleen cells were therefore carried out. After removal of macrophages by passage through Sephadex G-10 columns[11] the spleen cells failed to produce antibody to SRBCs. Responsiveness was restored by addition of 5 × 10⁻⁵ M 2-mercaptoethanol (2-ME)[12], but not by addition of TNS. In Sephadex-treated cultures supplemented with 2-ME, addition of TNS enhanced the generation of antibody-forming cells (Fig. 1). As the *in vitro* response to SRBCs requires the participation of helper T cells, experiments were carried out to determine the effect of TNS on T-cell-depleted populations. Spleen cells which had been depleted of T cells by treatment with anti-Thy-1.1 serum and complement (after passage through Sephadex G-10 columns) produced only few plaque-forming cells (PFCs) in culture; the production of antibody to SRBCs by the macrophage
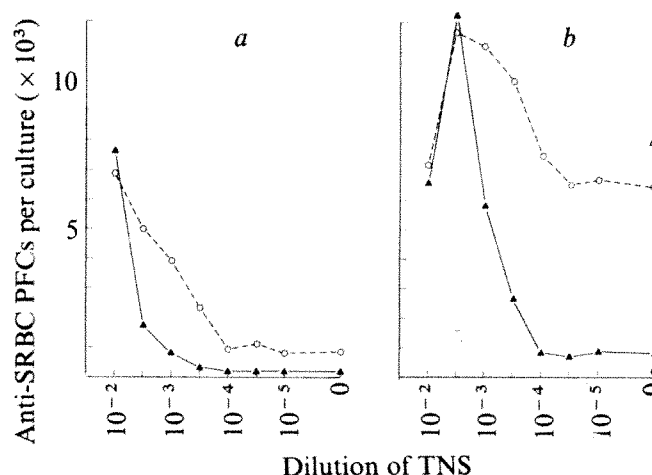
Fig. 1 Restoration of antibody production by macrophage-, and T-cell-depleted spleen cell cultures with TNS. PFC response on days 3(*a*) and 4(*b*). Each point is the average of 2 cultures. ●, 2-ME not added; ○, 2-ME added; ▲, 2-ME added, T cells removed by treatment with anti-Thy-1.1 and rabbit complement; △, 2-ME added, cells treated with rabbit complement (control).

and T-cell-depleted spleen cell population was restored by addition of TNS to the culture medium (Fig. 1).

Figure 2 shows results of experiments with congenitally athymic Nu/Nu mice. Since they lack helper T cells, spleen cells from these mice do not produce antibody to red blood cells *in vitro*. Addition of TNS to the culture medium permitted a response (of the same magnitude as that of spleen cells from littermate Nu/+ mice which do not lack T cells) to the antigen that was used to sensitise the cultures (burro red blood cells, BRBCs), but not to an unrelated antigen (SRBC). Addition of TNS was effective even if delayed as long as 2 d (Table 2), and this was also true for TNF. Thus, injection of LPS into BCG-infected mice causes release of a factor which can substitute for helper T cells in the production of antibody to heterologous red-cell antigen *in vitro*, an activity which

Fig. 2 Restoration of antibody production by spleen cells from Nu/Nu mice with TNS. Cultures (in MEM + 5% FCS) were immunised with BRBC. Anti-BRBC (*a*) and anti-SRBC PFC (*b*) were assayed on day 4. △, TNS; ●, normal mouse serum; ○, serum from mice treated with BCG only; □, serum from mice treated with LPS only.

Table 1 Effects of TNS on production of antibody to SRBCs in vitro

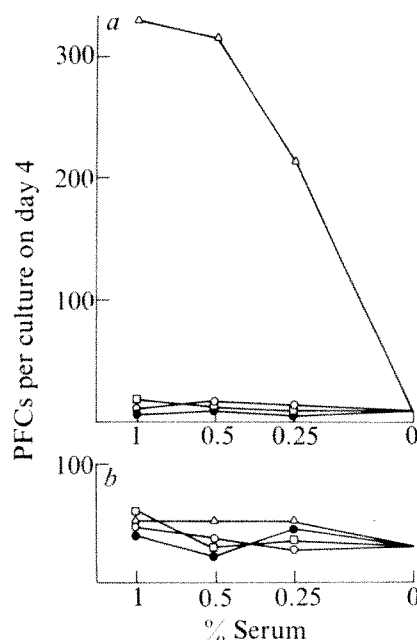| Addition to culture*	Anti-SRBC PFCs per culture	
	Day 3	Day 4
None	65	3,900
TNS	1,850	13,500
TNF	1,280	10,200
Normal mouse serum	28	4,000
LPS serum†	108	6,500
BCG serum‡	78	4,000

*Final concentration of serum 2%, TNF 0 13 mg ml^{-1}
†Serum from mice given LPS only.
‡Serum from mice given BCG only

Fig. 3 Induction of antibody production in cultures of CR⁻ spleen cells by TNS. Spleen cells were passed through Sephadex G-10 columns coated with complement-reacted antigen-antibody complexes (CR⁻ spleen cells, hatched bars) or uncoated Sephadex G-10 columns (CR⁺ spleen cells, open bars) TNS was added at 1% final concentration As source of macrophages (Mφ), 5 × 10⁴ peritoneal cells were added to the cultures. Average of 2 cultures

resides in the same serum fraction (TNF) as the tumour-necrotising activity

We have already indicated that TNS does not substitute for that function of macrophages which can be replaced by 2-ME. We have described another function of macrophages, namely, the support of antibody formation in vitro by mouse spleen cells which are depleted of B cells carrying complement receptors (CR⁺)[13]. CR⁻ spleen cells are incapable of cooperating with primed helper cells in culture unless macrophages are present. In this context, macrophages cannot be replaced by 2-ME, but they can be replaced by culture supernatants of peritoneal macrophages that have been activated in vivo by injection of BCG. We have now found that TNS induces CR⁻ spleen cells to produce antibody to SRBCs and to cooperate with primed helper cells.

CR⁻ spleen cells were obtained by passage of mouse spleen cells through Sephadex G-10 columns coated with complement-reacted antigen–antibody complexes. This type of column removes macrophages and CR⁺ B cells[13]. As a control, spleen cells were passed through uncoated Sephadex G-10 columns, which remove macrophages[11], but not CR⁺ B cells[13]. CR⁻ spleen cells from mice primed with SRBC were used as a source of pri,ed helper T cells[14]. One part primed cells (depleted of CR⁺ cells) was mixed with 9 parts unprimed CR⁻ cells, the cultures were sensitised with trinitrophenyl (TNP)-conjugated SRBCs, and the number of anti-TNP PFCs was determined on day 4. As shown in Fig 3, CR⁻ spleen cells not only generated considerably less anti-TNP PFCs than CR⁺ spleen cells, but also failed to cooperate with SRBC-primed helper T cells. TNS induces CR⁻ spleen cells to produce antibody and to cooperate with primed helper T cells in the absence of macrophages. We have postulated[13] that CR⁻ B cells must differentiate into CR⁺ B cells to carry out this function. This step in B-cell differentiation may be induced by macrophages when they are activated as, for example, after injection of BCG or by activated helper T cells[13]. As we will report elsewhere, TNF induces differentiation of CR⁻ B cells to CR⁺ B cells in vitro (M.H. and U Hammerling, unpublished).

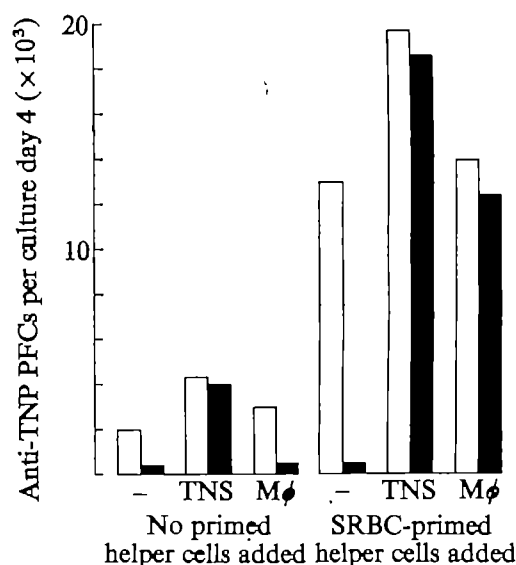The cellular origin of TNF is not as yet known. For various

reasons, macrophages have been considered a likely source[8,9]. The most striking phenomenon which occurs during the priming phase for TNF production is proliferation and activation of macrophages. Moreover, activated macrophages or culture supernatants of activated macrophages have been reported to possess activities that are similar to those of TNF. Like TNF they induce antibody formation in cultures of spleen cells from Nu/Nu mice, particularly after exposure to LPS. In addition, macrophages exposed to LPS become cytotoxic for transformed target cells in tissue culture[15,16], and TNF shares this property[8] Some effects of TNF, on the other hand, mimic effects of factors assumed to be derived from T cells. As these T-cell factors have been reported to react with antisera directed against Ia and β_2 microglobulin[5], use of such antisera may be useful for further definition of TNF. Until TNF has been fully characterised, however, it would be premature to ascribe the immunological effects reported here to the anti-tumour factor

M. K. HOFFMANN
S GREEN
L. J. OLD
H. F. OETTGEN

Memorial Sloan–Kettering Cancer Center,
1275 York Avenue,
New York, New York 10021

1 Huang, T., Schimpl, A., and Wecker, E., J exp Med, 139, 754–760 (1974)
2 Dutton, R W, Transplant Rev, 23, 66–77 (1975)
3 Schimpl, A L, and Wecker, E, Nature, 237, 15–17 (1972)
4 Dutton, R W, and Hunter, P, in Cellular Selection and Regulation in the Immune Response (edit by Edelman, G M), 199 (Raven, New York, 1974)
5 Armerding, D, Kubo, R, Grey, H, and Katz, D, Proc natn Acad Sci U.S.A, 72, 4577–4581 (1975)
6 Schrader, J W, J exp Med, 138, 1466–1480 (1973)
7 Watson, J, Trenkner, E, and Cohn, M, J exp Med, 138, 699–714 (1973)
8 Carswell, E A, et al, Proc natn Acad Sci U S A, 72, 3666–3670 (1975)
9 Green, S, et al, Proc natn Acad Sci U S A, 73, 381–385 (1976)
10 Mishell, R I, and Dutton, R W, J exp Med, 126, 423–442 (1967)
11 Ly, I A, and Mishell, R I, J Immun Meth, 5, 289–294 (1974)
12 Chang-Chen, M B, and Hirsch, J A, Science, 176, 60–61 (1972)
13 Hoffmann, M K, Hammerling, U, Simon, M, and Oettgen, H F, J Immun, 116, 1447–1451 (1976)
14 Hoffmann, M, and Kappler, J W, J exp Med, 137, 721–739 (1973)
15 Alexander, P, and Evans, R, Nature new Biol, 232, 76–78 (1971)
16 Currie, G A, and Basham, C, J exp Med, 142, 1600–1605 (1975)

Table 2 Substitution of T-cell helper function by simultaneous as well as delayed addition of TNS or TNF

Additions	Anti-SRBC PFCs per culture on day 4	
	(addition on day 0)	(addition on day 2)
None	<5	<5
TNS	170	390
TNF	230	310

Cultures of spleen cells of Nu/Nu mice were immunised with SRBCs on day 0 and assayed for anti-SRBC PFCs on day 4 TNS 1%, or TNF 0 13 mg ml⁻¹, were added as indicated on day 0 or day 2. Each value is the average of 3 cultures.

Dissociation of alcohol tolerance and dependence

CHRONIC.intoxication with ethanol has been shown to result in tolerance to and physical dependence on this drug[1,2]. In general, the development of tolerance and physical dependence during intake of addictive compounds have been looked on as inexorably linked phenomena produced by a common biochemical[3] or physiological[4] mechanism. This was supported by the studies of Way et al.[5] who demonstrated a simultaneous development and disappearance of tolerance and physical dependence after morphine pellet implantation in mice. We were interested in the possible application of the same theoretical framework to the development of tolerance to and physical dependence on ethanol. Previous work[6,7] demonstrating changes in noradrenergic systems in brain after chronic treatment led us to study the effect of 6-hydroxydopamine (6-OHDA) pretreatment on the development of tolerance to and physical dependence on ethanol.

In our studies, one group of C57BL/6 male mice was injected intracerebroventricularly[8] with 6-OHDA (50 µg in 10 µl artificial cerebrospinal fluid (CSF)[9] containing 5 mg ml^{-1} ascorbate), a second group of mice was injected with only the artificial CSF with ascorbate, and a third group of animals received artificial CSF containing no ascorbate. Since the presence of ascorbate in the artificial CSF had no effect on any of the parameters discussed below, all reported comparisons were between groups of animals injected with the CSF containing ascorbate and those injected with CSF, ascorbate and 6-OHDA. Brain levels of dopamine[10], noradrenaline[10] and 5-hydroxytryptamine (5-HT)[11] were determined in aliquots of two pooled brains[8,12] seven days after the intracerebroventricular injection and then again after chronic ethanol treatment. These levels were 6.47 ± 0.87, 1.47 ± 0.29 and 3.24 ± 0.48 nmol per g brain, respectively for the animals injected with 6-OHDA and 8.88 ± 1.60, 3.26 ± 0.46 and 3.44 ± 0.29 nmol per g brain, respectively for those injected with artificial CSF. The levels of these neurotransmitters were not significantly altered by the chronic ethanol treatment described below.

A week after the injection, mice were placed in individual cages and offered a liquid diet consisting of Carnation Slender, sucrose (100 g l^{-1}) and Vitamin Diet Fortification Mixture (ICN Pharmaceuticals) (3 g l^{-1}) as their only food source. Twenty-four hours after being placed on the liquid diet, mice were divided into four groups:

Group 1: animals injected with artificial CSF. This group continued to receive the sucrose-containing diet throughout the period when the other animals (groups 3 and 4) received ethanol.

Group 2: mice injected with 6-OHDA and continued on the diet containing sucrose as were those in group 1.

Group 3: mice injected with artificial CSF. These mice were fed a diet in which the sucrose was substituted with 7% v/v ethanol.

Group 4: mice injected with 6-OHDA and fed the diet containing the 7% v/v ethanol as were those in group 3. Animals were maintained on their respective diets for seven full days, and on the morning of day 8, all animals again received the diet containing sucrose (withdrawal). During the period that mice received ethanol, body weight, behavioural intoxication, body (rectal) temperature, quantity of diet consumed and blood ethanol levels[13] were monitored daily[14]. The quantity of diet given to the mice consuming the sucrose-containing mixture was made equal to the amount of ethanol-containing diet consumed by the animals in groups 3 and 4.

No differences were found in the daily amounts of diet consumed or blood ethanol levels during the chronic treatment with ethanol between animals in group 3 compared with those in group 4 (t test, $P > 0.25$). Behavioural intoxication scores[14] were also not significantly different between these two groups

of animals. Body temperature was, however, found to be lower in animals of group 4 compared with those of group 3 (t test, $P < 0.05$) during the last two days on the ethanol diet. A more complete description of the animals during the period of chronic ethanol intake will appear elsewhere (Ritzmann and Tabakoff, in preparation).

After withdrawal from ethanol, animals were observed over a 12-h period: six 1-h periods of observation alternating with 1-h periods during which no observation took place. During the observation periods, body temperature was measured and behavioural signs of withdrawal were noted[14]. Withdrawal signs were scored as follows: 0 = no withdrawal symptoms; 1 = characteristic tremors and spasms when mice are lifted by the tail[15]; 2 = spontaneous tremors and spasms; 3 = spontaneous seizures (clonic and/or tonic)[14]; and 4 = death.

The behavioural withdrawal scores and measures of the concomitant hypothermia[14] are summarised in Table 1. No significant differences (see legend to Table 1) in either the severity of the hypothermia or the behavioural manifestations of withdrawal were evident when animals from groups 3 and 4 were compared. Animals receiving the sucrose-containing
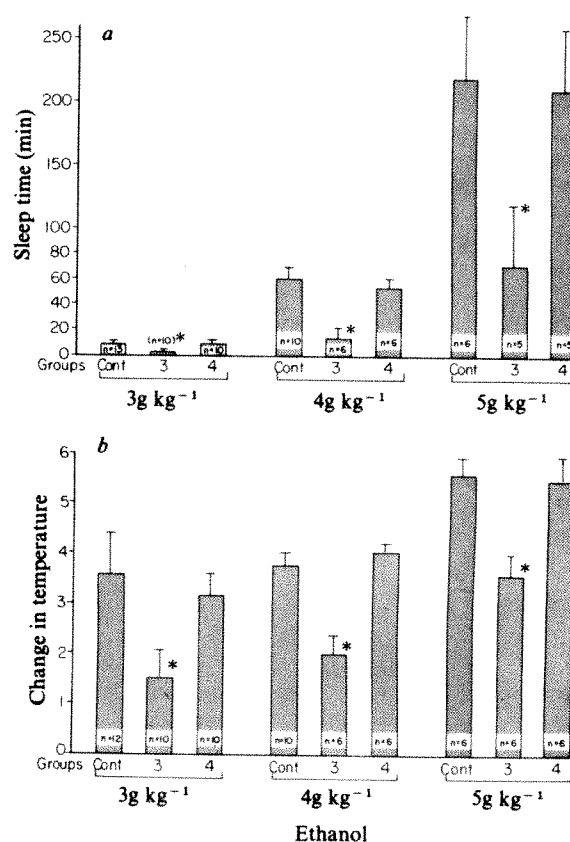
Fig. 1 Body temperature and sleep time after ethanol. Mice were injected intraperitoneally with ethanol (3, 4 and 5 g kg^{-1}) 26 h after withdrawal, and sleep time (a) and body temperature (b) were monitored. Since no differences in these parameters or blood ethanol levels were found between animals in groups 1 and 2, the values obtained with these animals were pooled and are labelled "control." Sleep time represents the time from loss of righting reflex to the time the mouse could right itself twice in 30 s. The bars represent the mean$\pm$s.d., and the number within each bar is the number of animals used for the determination. Rectal temperature was monitored at 15- and 30-min intervals for 3 h after ethanol injection. The bars represent the mean$\pm$s.d. of the greatest change from preinjection temperatures recorded during this period. The nadir for the 3 and 4 g kg^{-1} dose occurred at 30 min and for the 5 g kg^{-1} dose at 60 min after injection. Comparison between groups at each dose was made using the t test as well as the Mann–Whitney U test. *Groups significantly different from each of the other groups ($P < 0.05$) at a particular dose; other groups were not significantly ($P > 0.2$) different from one another.

Table 1 Hypothermia and withdrawal symptoms in mice

Behavioural withdrawal score	Time after withdrawal (h)					
	0	2	4	6	8	10
Group 3:						
0	36.2±0.4 (n = 10)	36.±0.4 6 (n = 9)	36.1±0.1 (n = 2)	—	36.9±0.2 (n = 4)	36.9±0.4 (n = 11)
1	—	35.0 (n = 1)	35.1±0.7 (n = 6)	35.8±0.6 (n = 6)	35.9±0.6 (n = 6)	36.0 (n = 1)
2	34.0 (n = 1)	34.2 (n = 1)	34.6±0.3 (n = 4)	34.2±0.7 (n = 6)	35.7±0.1 (n = 2)	—
3	33.0 (n = 1)	33.5 (n = 1)	—	—	—	—
4	—	—	—	—	—	—
Group 4:						
0	35.3±0.4 (n = 12)	36.5±0.5 (n = 11)	36.4 (n = 1)	36.4 (n = 1)	36.9±0.1 (n = 2)	36.6±0.5 (n = 8)
1	—	35.8 (n = 1)	35.1±0.7 (n = 6)	35.5±0.5 (n = 5)	36.3±0.9 (n = 4)	35.8±1.2 (n = 3)
2	—	—	34.3±0.6 (n = 4)	34.4±0.4 (n = 4)	35.1±0.9 (n = 5)	—
3	—	—	—	34.8 (n = 1)	—	—
4	—	—	(n = 1)	—	—	—

Ethanol was removed from the diet of mice after 7 d of chronic consumption, and behaviour and body temperature were monitored. Mice receiving similar withdrawal severity scores at the various times of testing were grouped and the mean ± s.d. of the body temperature for each such group was determined and noted in the table. n represents both the number of animals exhibiting the particular withdrawal symptoms at the time of testing and the number of animals used to determine the mean temperature for the group. Hypothermia occurring in groups 3 and 4 was compared, after pooling the temperature values at each time point, using Student's t test. To assess whether or not behavioural withdrawal scores were different between groups 3 and 4, mice were subdivided as those receiving a score of 0 or 1 and those receiving a score of 2 or higher at each time point. Statistical analysis was performed using the Fischer Exact Probability test. No statistically significant differences were found between the body temperatures or withdrawal scores of groups 3 and 4.

control diet (groups 1 and 2) demonstrated none of the withdrawal symptoms. In separate experiments withdrawal hyperexcitability was assessed by injecting animals with pentylenetetrazol (60 mg kg^{-1}, intraperitoneally) and monitoring the incidence of convulsive seizures and death occurring within a 30-min period following such injection[16]. Pentylenetetrazol was administered to animals between 6 and 7 h after withdrawal of ethanol.

Animals which had been withdrawn from the ethanol-containing diets were significantly (binomial test[17]; P in all cases < 0.025) more susceptible to pentylenetetrazol-induced convulsions than sucrose-fed controls. No statistically significant differences (P > 0.3) in severity or frequency of pentylenetetrazol-induced seizures were, however, noted when comparing animals of group 3 with those of group 4. Eight animals from each of the four groups were tested, and the number of animals exhibiting clonic or tonic convulsions or dying in seizures after treatment with pentylenetetrazol in group 1 was two; in group 2, two; in group 3, six, and in group 4, seven.

Twenty-six hours after withdrawal, when all overt signs of the withdrawal syndrome were no longer present, mice were injected intraperitoneally with ethanol and "sleep time"[18] and rectal temperature were monitored. Mice previously injected with pentylenetetrazol were not used in these studies.

Animals of group 3 were found to be tolerant to the effects of ethanol by measures of both rectal temperature and "sleep time" (Fig. 1). On the other hand, pretreatment of the animals with 6-OHDA before the chronic ethanol treatment (group 4) resulted in a block in development of tolerance to the measured effects of ethanol (Fig. 1). No significant differences in blood ethanol levels during the time that behavioural and physiological parameters were being measured were evident between any of the groups after injection of ethanol (3 g kg^{-1}).

Thus, animals treated with 6-OHDA and force fed a diet containing ethanol were found to become physically dependent on ethanol (as demonstrated by the appearance of withdrawal symptoms), but did not become tolerant to two of the major physiological effects of ethanol administration, that is, hypothermia and sedation.

Administration of 6-OHDA has been shown to result in the

destruction of catecholamine-containing neuronal terminals[19]. Such destruction of the functional connections between the noradrenergic neurones and other brain neurones seems to prevent the necessary alterations in neural cybernetics which produce tolerance after chronic ethanol treatment. Although the injections of 6-OHDA did not produce a statistically significant (t = 1.87, d.f. = 6) decrease in the levels of dopamine in our mice, the levels were substantially altered. Thus, we do not feel that the block in the development of tolerance can at present be totally ascribed to the destruction of only the noradrenergic systems.

The dissociation of tolerance and physical dependence witnessed in the present study provides an important conceptual framework within which the neurological effects of ethanol should be viewed. Seevers and Woods[20] and Kissin[21] suggested previously that tolerance and physical dependence may not necessarily be unitary phenomena, and our studies demonstrate that certain manifestations of tolerance to ethanol may be eliminated without affecting the development of physical dependence produced by this drug.

RONALD F. RITZMANN
BORIS TABAKOFF

Department of Physiology,
University of Illinois Medical Center,
Chicago, Illinois 60612

1 Mendelson, J. H., New Engl. J. Med., 283, 24–81 (1970).
2 Kalant, H., LeBlanc, A. E., and Gibbins, R. J., Pharmac. Rev., 23, 135–172 (1971).
3 Goldstein, A., and Goldstein, D. B., in The Addictive States (edit. by Wikler, A.), 265–267 (Williams–Wilkins, Baltimore, 1968).
4 Jaffe, J. H., and Sharpless, S. K., ibid., 226–243.
5 Way, E. L., Loh, H. H., and Shen, F. H., J. Pharmac. exp. Ther., 167, 1–8 (1969).
6 Pohorecky, L. A., J. Pharmac. exp. Ther., 189, 380–391 (1974).
7 Hunt, W. A., and Majchrowicz, E., J. Neurochem., 23, 549–552 (1974).

8 Tabakoff, B., Ritzmann, R. F., and Boggan, W. O., *J. Neurochem.*, **24**, 1043–1051 (1975).
9 Merlis, J. K., *Am. J. Physiol.*, **131**, 67–72 (1940–41).
10 Anton, A. H., and Sayre, D. F., *J. Pharmac. exp. Ther.*, **138**, 360–375 (1962).
11 Maickel, R. P., Cox, R. H., Jr, Saillant, J., and Miller, F. P., *Int. J. Neuropharmac.*, **7**, 275–281 (1968).
12 Neff, N. H., Spano, P. F., Groppetti, A., Wang, C. T., and Costa, E., *J. Pharmac. exp. Ther.*, **176**, 701–710 (1971).
13 Tabakoff, B., Anderson, R. A., and Ritzmann, R. F., *Biochem. Pharmac.*, **25** (11), 1305–1309 (1976).
14 Ritzmann, R. F., and Tabakoff, B., *J. Pharmac. exp. Ther.* (in the press).
15 Goldstein, D. B., *J. Pharmac. exp. Ther.*, **183**, 14–22 (1972).
16 Tabakoff, B., and Boggan, W. O., *J. Neurochem.*, **22**, 759–764 (1974).
17 Siegel, S., *Nonparametric Statistics for the Behavioural Sciences*, 36–42 (McGraw-Hill, New York, 1956).
18 Tabakoff, B., thesis, Univ. Colorado, Boulder (1970).
19 Sachs, C., and Jonsson, G., *Biochem. Pharmac.*, **24**, 1–8 (1975).
20 Seevers, M. H., and Woods, L. A., *Am. J. Med.*, **14**, 546–557 (1953).
21 Kissin, B., in *Biology of Alcoholism*, 3 (edit. by Kissin, B., and Begleiter, H.), 1–36 (Plenum, New York, 1974).

Progressive glomerulonephritis in mice treated with interferon preparations at birth

IN addition to the well known antiviral action[1] interferon preparations exert various biological effects on cells (for references see refs 2 and 3). While investigating the effect of interferon on the growth and development of newborn mice, we found that daily injection of potent mouse interferon preparations resulted in extensive liver cell degeneration and death between days 8 and 14 (ref. 4). When interferon treatment, begun at birth, was stopped between days 6 and 8 of life, liver cell damage appeared to be reversible and most mice recovered. In the ensuing months, several of these mice died and although the liver and other organs appeared normal, the kidneys were pale and the surface was granular. Histological examination revealed a severe glomerulonephritis. We report here an investigation of this phenomenon.

Newborn Swiss (and in one experiment C3H mice) from specific pathogen-free colonies were injected subcutaneously in the interscapular region with 0.05 ml of mouse interferon or the different control preparations daily from day 0 to days 6–8, at which time treatment was stopped. When interferon-treated mice were killed on day 8, histological sections of liver showed focal cellular necrosis, as reported before[4], but the kidneys appeared normal. Nevertheless, Swiss mice began to die with renal lesions as early as day 35. Autopsy of some mice dying with kidney disease showed signs suggestive of cardiac failure, that is enlarged heart, and haemorrhagic and oedematous lungs.

Mice were killed at regular intervals, and sections of organs were examined under light microscopy. In all cases, the kidneys of Swiss mice treated with interferon at birth showed glomerulonephritis (Table 1) whereas the liver appeared normal. The kidney pathology can be summarised as follows. In "early" lesions (Swiss mice killed on the 20th day of life, Table 1) the glomeruli appeared moderately enlarged with some thickening of the mesangium, some segmental foci of mesangial proliferation and some thickening of peripheral capillary loops. There were no tubulo-interstitial lesions except for a few mononuclear cells. The vessels appeared normal. By immunofluorescence there were focal and segmental granular deposits of IgG, IgM and C3 in the mesangial spaces and along some glomerular basement membranes (GBM). Fibrinogen and IgA were not detected. In Swiss mice killed after 31 d, the lesions were more conspicuous with hyalinisation of glomerular tufts (with epithelial crescents and voluminous subendothelial deposits) (Fig. 1a). In well advanced disease, virtually all glomeruli were sclerotic. Amyloid was not detected. The tubulo-interstitial lesions paralleled the glomerular damage, showing extensive sclerosis, and diffuse atrophy of tubules with dilated lumina filled with hyalin and proteinaceous casts (Fig. 1b). By immuno-

Table 1 Appearance of glomerulonephritis in mice treated for 6–8 d from birth with mouse interferon preparations

Strain of mouse	No. of experiments	Age of mice killed (d)	Mouse* interferon	Control preparation	Inactivated† mouse interferon	Human‡ interferon	Uninoculated
		20	4/4 §(1.2)				0/2
		31–35	10/10 (1.9)	0/3	0/3	0/3	0/5
		45–60	7/7 (3.3)	0/9			0/7
		67–82	12/12 (3.3)	0/2	0/4	0/4	0/5
Swiss	9	85–100	8/8 (3.4)	0/1			0/5
		102–124	10/10 (2.8)	0/8	0/5	0/5	0/11
		143–185	9/9 (2.4)	0/5	0/6	1/6¶	0/7
		198–274		0/22			2/23‖
	Total		60/60	0/50	0/18	1/18	2/65
C3H	1	43–55	8/8 (2.0)				0/5
		104–132	5/9 (0.6)				0/8
	Total		13/17				0/13

Mouse interferon was prepared from Swiss mouse C-243 cells induced with Newcastle disease virus (NDV), and concentrated, partially purified and assayed as decribed before[5]. Mock interferon was prepared in a manner identical to that used in the preparation of interferon with the exception that the interferon inducer, NDV, was either omitted, or added for 1 h immediately before collection of the culture supernatant. All interferon and control preparations were dialysed further for 24 h at 4 °C against 0.01 M perchloric acid before testing for toxicity on a line of L 1210 cells resistant to interferon[6]. The specific activity of mouse C-243 cell interferon preparations was in the range 0.7–3.0 × 10⁶ U per mg protein. In one experiment mice were treated with a more purified C-243 interferon preparation (1.9 × 10⁷ interferon reference U per mg protein) (purified by affinity chromatography using sheep anti-mouse interferon globulins). C-243 cell control preparations contained approximately the same amount of protein per ml as the C-243 cell interferon (that is 1.5–3.4 mg ml⁻¹). Partially purified human interferon (specific activity 3.5 × 10⁵ U per mg of protein) prepared in leukocyte suspensions inoculated with Sendai virus was provided by Dr K. Cantell[7]. Mouse interferon units quoted are as measured in our laboratory and one of our units equals 4 mouse interferon reference units.
* A different mouse C-243 cell interferon preparation (titre 1:800,000 to 1:1,600,000) was used in each experiment. Only mice having been injected daily from birth to days 6 to 8 of life are included.
† Mouse C-243 interferon with a titre of 1:800,000 was heated at 60 °C for 2 h at pH 10. Residual titre was 1:20.
‡ Titre of human leukocyte interferon was 1:1,000,000 (MRC reference interferon 69/19 U) as assayed on human diploid fibroblasts.
§ No. of mice with histological confirmation of glomerulonephritis per total number of mice killed. Each kidney was examined by light microscopy and the severity of lesions was graded on a scale from 0 to 4. In all cases there was little variation in the extent of lesions within a given series and the figures in brackets indicate the mean index of severity of kidney lesions for each series.
¶ The kidney of one mouse showed glomerulonephritis grade 2.
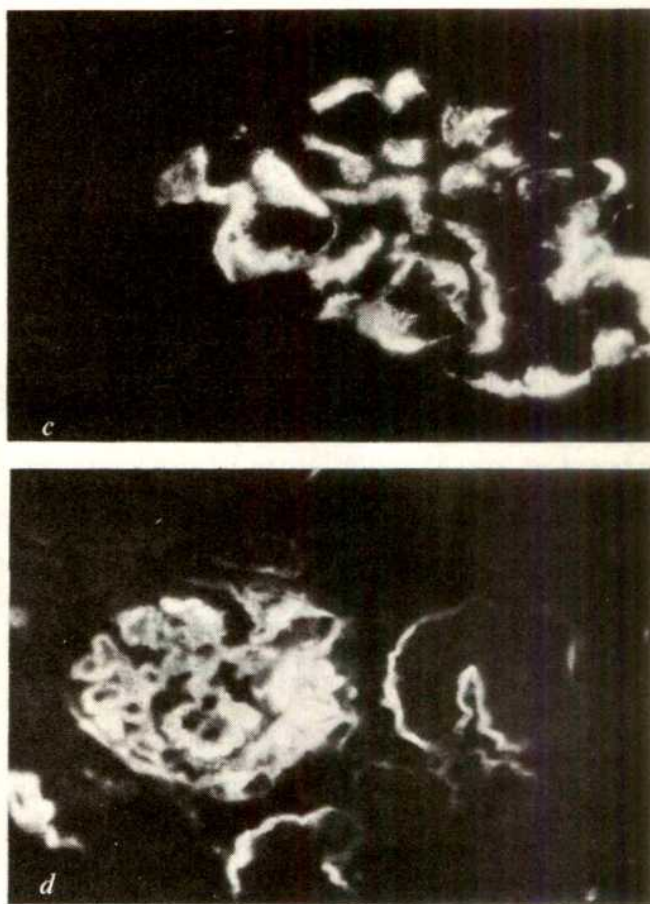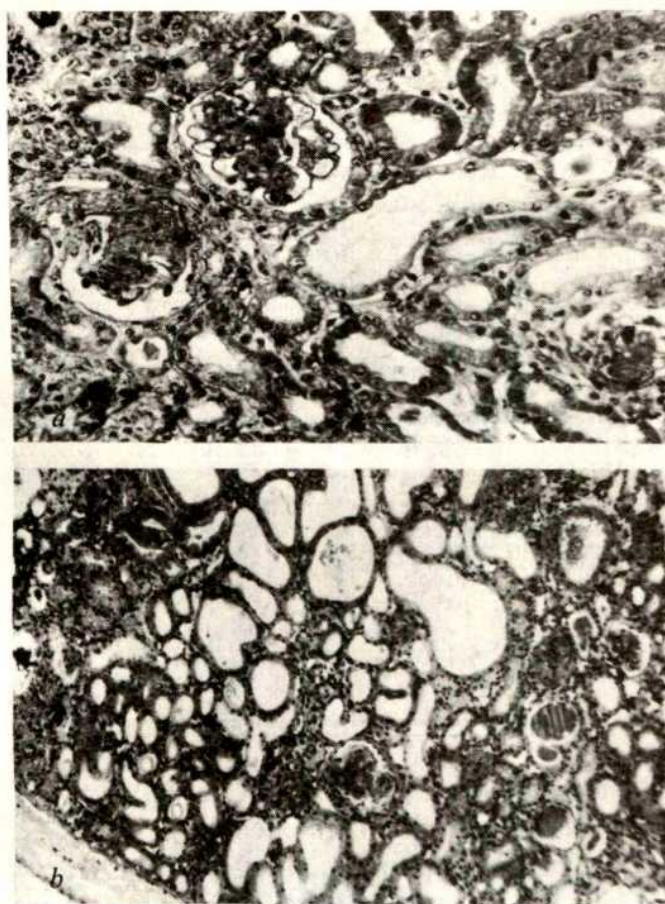‖ The kidneys of two mice showed glomerulonephritis grade 1.

Fig. 1 Kidneys of Swiss mice treated with mouse interferon for the first 7–8 d of life. *a*, Mouse killed at 100 d. Note mild glomerular proliferation and hyalinisation of glomerular tufts. Tubular epithelium appears flattened with dilatation of lumina. There are a few inflammatory cells in the interstitium (×250). *b*, Mouse killed at 60 d. Note extensive sclerosis of interstitial tissue (with cellular infiltration) with pronounced atrophy of tubules and sclerotic glomeruli (×100). *c*, Mouse killed at 85 d. Stained with fluoresceinated anti-mouse IgG serum. Note coarse granular deposits along GBM (×540). *d*, Stained with fluoresceinated anti-mouse C3 serum. Note deposits of C3 along GBM. The deposits of C3 along the tubular basement membrane have no pathological significance as they are also present in control mice (×400). Renal tissue was fixed in Dubosq–Brazil fixative (alcoholic Bouin) for 2 d and then in 4% formalin in saline, and cut at 3 μm. The following stains were used systematically: Mallory's trichromic, periodic acid-Schiff, haematoxylin and eosin, Weigert's fuchsin, Wilder's reticulin and crystal violet. Immunofluorescence was done on snap-frozen specimens cut in a cryostat. Rabbit anti-mouse IgG, IgA, IgM, C3 and fibrinogen sera were obtained from Cappell Laboratories. The specificity of each antiserum was checked by immunoelectrophoresis. All magnifications are original magnifications.

fluorescence there were granular coarse deposits of IgG, IgM and C3 along the GBM (Fig. 1*c* and *d*). Fibrinogen was detected in crescents. IgA was not detected.

The results of our experiments can be summarised as follows. (1) In nine experiments, all Swiss (60 out of 60) mice treated with potent interferon preparations for the first 6–8 d of life developed severe glomerulonephritis early in life (Table 1). Only one out of 86 mice treated with the various control preparations and two out of 65 untreated mice developed a moderate glomerulonephritis and this occurred only in mice kept for more than 143 d (Table 1). (The finding of glomerulonephritis in some ageing Swiss mice is in accord with the results of Oldstone and Dixon who reported a "mild nephritis seen in 7% of SWR/J control mice by 9 months"[8].)

(2) Glomerulonephritis developed only in mice inoculated with enough interferon and for a sufficient time. Thus, only mice inoculated with interferon preparations titering approximately 1 : 10^6 for 6–8 d developed glomerulonephritis. Mice injected for 8 d with interferon preparations of lower titre, that is 1 : 10^5 (23 mice), or 1 : 10^4 (18 mice), showed no signs of glomerulonephritis when killed at 85, 140 and 198 d. In addition, injection of potent interferon preparations (1 : 10^6) did not result in glomerulonephritis when injected for only the first 2 d of life (20 mice killed at 85, 140 and 198 d).

(3) Interferon treatment resulted in glomerulonephritis only when administered during the first week of life. Thus, none of nine mice inoculated between the 8th and 15th d of life with an interferon preparation titering 1 : 10^6 showed glomerulonephritis when killed at day 59. (4) Male and female mice seemed to be equally susceptible as judged by the severity of disease and the incidence of death.

(5) The progression of the disease in Swiss mice was observed by killing mice at intervals and by following the progression of lesions in the same mouse after unilateral nephrectomy. Kidney lesions were minimal at 3 weeks, definite at 1 month and considerably more advanced in the ensuing weeks. (Table 1: note figures in brackets indicating mean index of severity of kidney lesions graded on a scale from 0 to 4). A kidney from each of two interferon-treated mice was removed on day 31 and the mice were killed on day 45 and 60. In both cases a progression of lesions in the remaining kidney was observed. (Unilateral nephrectomy in control mice did not cause glomerulonephritis in the contralateral kidney.) Since a number of interferon-treated mice died (probably of renal failure) after 4 weeks of life, it is possible that the apparent decrease in the severity of kidney lesions in mice older than 102 d (Table 1) reflects the selection of surviving mice with less severe disease rather than a reversibility of the pathological process.

(6) Although glomerulonephritis also developed in C3H mice treated from birth for 6–8 d with mouse interferon preparations titering 1 : 10^6 (according to the protocol used for Swiss mice) (Table 1), the glomerular lesions appeared

to be less severe than in Swiss mice at the same age, and deaths were not observed. Kidneys of eight out of eight C3H mice killed at 43–55 d showed moderate glomerulonephritis (Table 1). At 104–132 d, although glomerular lesions were present in six out of nine mice, these were less marked than in mice killed earlier. Furthermore, a kidney biopsy, at 45 d, of an interferon-treated C3H mouse showed moderate disease, whereas only minimal lesions were present when this mouse was killed at 104 d. Thus the glomerulonephritis in C3H mice is not only less severe than in Swiss mice, but may be reversible. (It is interesting in this regard that C3H mice developed a less severe glomerulonephritis than SWR/J mice when infected transplacentally or neonatally with lymphocytic choriomeningitis virus[*].)

Because there was a lapse of several weeks between cessation of interferon treatment (days 6–8) and the appearance of glomerulonephritis (day 20), and as interferon is cleared rapidly from the circulation[10-13], it seems unlikely that the development of progressive nephritis was due to a local retention of interferon. It is possible, however, that interferon damaged the kidney during the first week of life directly, or indirectly perhaps through its toxic effect on the liver[4,14]. Alternatively, this treatment might have altered the maturation of the kidney or some other organ(s), such as the spleen[4], leading ultimately to the development of glomerulonephritis.

The rapidity of onset of the glomerulonephritis and the histological lesions themselves do not permit ready classification in terms of pathogenesis[15-19]. The absence of linear deposits of Ig argues against the hypothesis of an anti-GBM glomerulonephritis and anti-GBM antibodies were not present in the sera of sick mice tested by indirect immunofluorescence. The granular pattern of the Ig and C3 deposits seem to favour the hypothesis of an immune complex nephritis, although we have no idea about the nature of a hypothetical antigen. We have never detected neutralising antibody to interferon in either the serum or γ-globulin-containing eluates of the kidneys of interferon-treated mice (neither was interferon itself present in these eluates), nor is there any evidence to suggest that interferon treatment during the first week of life activates a latent virus[4]. Anti-DNA antibodies were also not detected in the serum of sick mice.

Several chronic virus infections in animals can result in the late development of nephritis which is associated with the progressive deposition of virus–antibody complexes in the glomerulus[20-22]. Perhaps some of these viruses induce interferon in the neonatal period, contributing to the further development of glomerulonephritis. Furthermore, if our results in mice can be extrapolated to man, they suggest that an acute virus infection of the human newborn, if associated with the production of sufficient interferon, might result later in life in the development of a progressive glomerulonephritis.

ION GRESSER
CHANTEL MAURY
MICHAEL TOVEY

Institute de Recherches,
Scientifiques sur le Cancer,
BP 8-Villejuif,
France

LILIANE MOREL-MAROGER
FRANCOISE PONTILLON

Unité INSERM Néphrologie Normale,
*et Pathologique U64,*
Hôpital Tenon,
Paris, France

1 Isaacs, A., and Lindenmann, J., *Proc. R. Soc.*, **B147**, 258–267 (1957).
2 Gresser, I., Brouty-Boyé, D., Thomas, M. T., and Macieira-Coelho, A., *Proc. natn. Acad. Sci. U.S.A.*, **66**, 1052–1058 (1970).
3 Lindahl, P., Leary, P., and Gresser, I., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 2785–2788 (1973).
4 Gresser, I., Tovey, M. G., Maury, C., and Chouroulinkov, I., *Nature*, **258**, 76–78 (1975).
5 Tovey, M., Begon-Lours, J., and Gresser, I., *Proc. Soc. exp. Biol. Med.*, **146**, 809–815 (1974).
6 Gresser, I., Bandu, M. T., and Brouty-Boyé, D., *J. natn. Cancer Inst.*, **52**, 553–559 (1974).
7 Cantell, K., Hirvonen, S., Mogensen, K. E., and Pyhälä, L., *In Vitro*, Monog. No. 3 (edit. by Waymouth, C.), 35–38 (Tissue Culture Ass., Rockville, Maryland, 1974).
8 Oldstone, M. B. A., and Dixon, F. J., *J. Immun.*, **106**, 1260–1263 (1971).
9 Oldstone, M. B. A., and Dixon, F. J., *J. Immun.*, **105**, 829–837 (1970).
10 Baron, S., Buckler, C. E., McCloskey, R. V., and Kirschstein, R. L., *J. Immun.*, **96**, 12–16 (1966).
11 Subrahmanyan, T. P., and Mims, C. A., *Br. J. exp. Path.*, **47**, 168–176 (1966).
12 Gresser, I., Fontaine, D., Coppey, J., Falcoff, R., and Falcoff, E., *Proc. Soc. exp. Biol. Med.*, **124**, 91–94 (1967).
13 Ho, M., in *Interferons and Interferon Inducers* (edit. by Finter, N.), 241–249 (North-Holland, Amsterdam and London; American Elsevier, New York, 1973).
14 Sakaguchi, H., Dachs, S., Mautner, W., and Grishman, E., *Lab. Invest.*, **13**, 1418–1426 (1964).
15 Dixon, F., *Harvey Lectures*, **58**, 21–52 (1963).
16 McCluskey, R. T., Vassalli, P., Gallo, G., and Baldwin, D. S., *New Engl. J. Med.*, **274**, 695–701 (1966).
17 McCluskey, R. T., in *Pathology of the kidney*, 1 (edit. by Heptinstall, R. H.), 273 (Little Brown, Boston, 1974).
18 Germuth, F. G., and Rodriguez, E., *Immunopathology of the Renal Glomerulus* (Little Brown, Boston, 1973).
19 Wilson, C. B., and Dixon, F. J., *A. Rev. Med.*, **25**, 83–98 (1974).
20 Tonietti, G., Oldstone, M. B. A., and Dixon, F. J., *J. exp. Med.*, **132**, 89–109 (1970).
21 Oldstone, M. B. A., and Dixon, F. J., *J. exp. Med.*, **134**, 32s–40s (1971).
22 Oldstone, M. B. A., *Prog. med. Virol.*, **19**, 84–119 (1975).

Postponement of symptoms of hereditary muscular dystrophy in chickens by 5-hydroxytryptamine antagonists

THE commoner forms of muscular dystrophy, both in man and in animals, are determined by a single gene change, but the mechanism whereby the complex set of symptoms develops is unknown. Chickens homozygous for muscular dystrophy have been bred, and form probably the closest animal model for the human disease so far investigated[1,2]. Dystrophic line 304 of the Davis (California) flock, for example, has been bred for rapid appearance of the symptoms and shows some signs of muscle weakness at 7–10 d *ex ovo*, whereas a normal line (200) of chickens with closely related genetic background is also available[1,2]. Two of us have reviewed[3] evidence that 5-hydroxytryptamine (5-HT) is involved in some way in the development of the symptoms of muscular dystrophy, and reported that administration of a 5-HT antagonist, methysergide, retards the development of the symptoms in line 304 chickens. We report here that this effect is more general, in that a second anti-serotoninergic drug of entirely different chemical struc-

Fig. 1 Mean FN (○) and plasma CPK level[8] (●) for untreated dystrophic chickens, as a function of age (50–82 animals were used per point). The horizontal line shows the constant FN for the normal chickens from day 10 onwards.
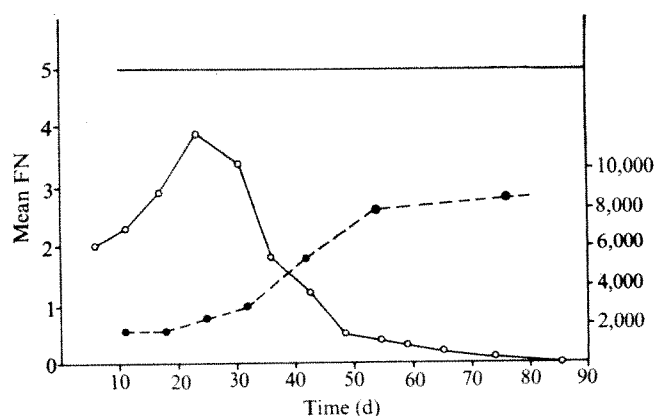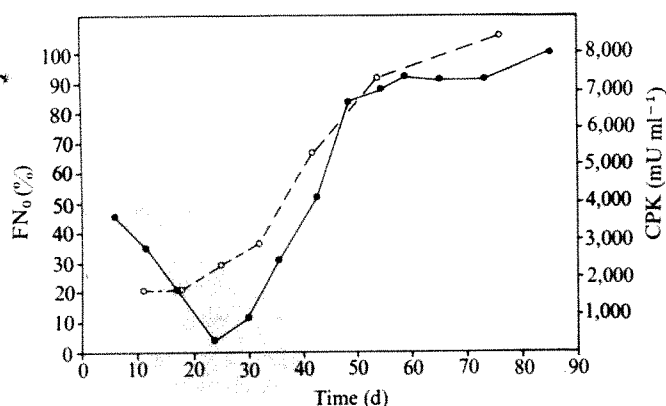
Fig. 2 Percentage of the untreated dystrophic population (50–82 animals per point) giving FN = 0 (●), and the corresponding plasma CPK level (○).

ture, cyproheptadine, is also effective, and that a combination of the two agents is beneficial.

We first established that the development of measureable symptoms in line 304 dystrophic chickens is regular enough for any drug-induced improvement to be reliably detected. A test[1,4,5] of muscle weakness in dystrophic chickens was used in the form of the 'flip number' (FN)—that is, the number of times (0–5 out of 5) a bird can right itself when placed on its back five times in immediate succession. The considerable increase in the creatine phosphokinase (CPK) content of the blood plasma, known to characterise this disease in the chicken[5,6] and in man[7], was used as a second criterion of muscle damage. Using a total population of 150 dystrophic chickens (line 304, Department of Avian Sciences, University of California, Davis) maintained in groups from different hatches over a 2-yr period, these two parameters were followed from hatching for 80 d. Chickens were maintained in constant conditions (temperature 30–33 °C, 12 h light/dark cycle), and the flip test was given in a precisely standardised manner by a naive observer, at 5-d intervals to avoid an exercise component. CPK testing was carried out in a completely blind fashion on coded sera in another laboratory. Newly hatched chicks of the normal line do not develop sufficient muscle strength to give FN=5/5 until 7–10 d of age and show high variability until then, but thereafter are constant, all at 5/5 (Fig. 1). The dystrophic chicks improve partially until about 3 weeks of age, and thereafter show a sharp decline to low scores, eventually zero. The plasma CPK level increases by about 400% throughout the latter period (Fig. 1). We have found that a very good correlation between the muscle weakness and release of CPK from the muscle can be demonstrated (Fig. 2) using as the measure of the former the FN_0 value—that is, the percentage of the dystrophic population that fails totally (having FN=0/5).

With this baseline for comparison, birds injected[3] (twice

daily, intraperitoneally) with cyproheptadine hydrochloride (Periactin, donated by Merck, Sharp and Dohme, Rahway, New Jersey) from day 3 *ex ovo*, were tested. Untreated animals were given water injections similarly throughout. At a low dose regime (for the chicken), substantial improvement in the FN value was shown (Fig. 3). A medium level dose (12 mg kg⁻¹ d⁻¹) extended the later period of this effect, but the performance still declined sharply after day 55. The latter decline was further postponed by a combination of higher doses of cyproheptadine and methysergide (Fig. 3). For the latter case, all the treated animals remained at a passing level (FN=4–5) up to day 60, whereas all of the untreated birds were below this level from about day 30. Similarly, the FN_0 values for animals treated with cypro-
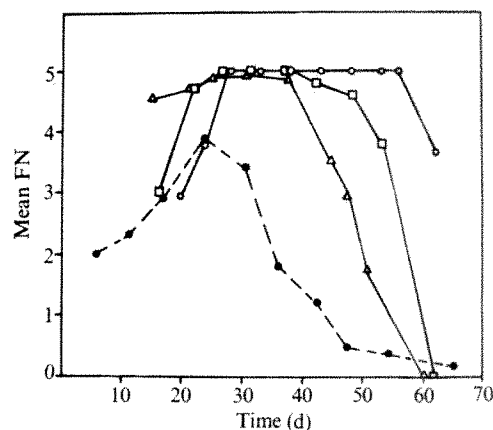

Fig. 3 Mean FN of cyproheptadine-treated or untreated (●), (see Fig. 1) dystrophic chickens. Low dosage regime (△, n = 12) was 2.4 mg kg⁻¹ d⁻¹ at days 3–16, 6.0 at days 17–43 and 8.0 at day 44 onwards. Medium dosage (□, n = 9) was 12.0 mg kg⁻¹ d⁻¹ from day 3 onwards. A combined treatment (○, n = 27) was also given, with cyproheptadine at 8.0 mg kg⁻¹ d⁻¹ at days 3–10, 12 at days 11–30, 17.5 at days 31–52, and 23 at day 53 onwards, plus methysergide at 12.5 mg kg⁻¹ d⁻¹ at days 3–15, 20 at days 10–30, 41 at days 31–45, and 52 at day 45 onwards.

heptadine or with the combination showed a pronounced difference from the untreated dystrophics (Fig. 4).

Using the criterion of CPK release, cyproheptadine-treated animals also showed a smaller change than the untreated birds (Fig. 5). The tendency to stabilise at a plateau of about 2,500 mU CPK per ml plasma in the treated birds, compared with about 8,500 in the untreated birds, is seen more clearly for the treatment with the two drugs combined.

Statistical data for points where divergences were

Table 1 FN differences in treated and untreated chickens

	10–13	Days *ex ovo* 41–44	56–60
Untreated	2.3±0.3 (57)	1.2±0.2 (56)	0.3±0.2 (56)
Methysergide	3.3±0.6 (15) [12.5]	5.0±0.0 (11)* [41]	4.0±0.7 (10)* [51]
Cyproheptadine	3.0±0.4 (9)	4.8±0.2 (9)*	1.7±0.8 (9)**
Methysergide + cyproheptadine		5.0±0.0 (27)*	4.6±0.1 (27)*

Values represent mean FN±s.e.m. for dystrophic chickens, with the number of animals used in parentheses. For the progressive dosages of cyproheptadine (medium dose) and the combination, see Fig. 3; for methysergide the dose level at each period is given in square brackets (in mg kg⁻¹ d⁻¹). Significance (using *t* test) from untreated dystrophic group: *$P < 0.001$; **$P < 0.01$.

Fig. 4 Percentage of the population of dystrophic chickens having FN = 0, when untreated (●, n = 15), or treated with cyproheptadine (○, dose low or medium, see Fig. 3; n = 21) or treated with both drugs (□, doses as stated for Fig. 3; n = 27, or, for last point only, 15).
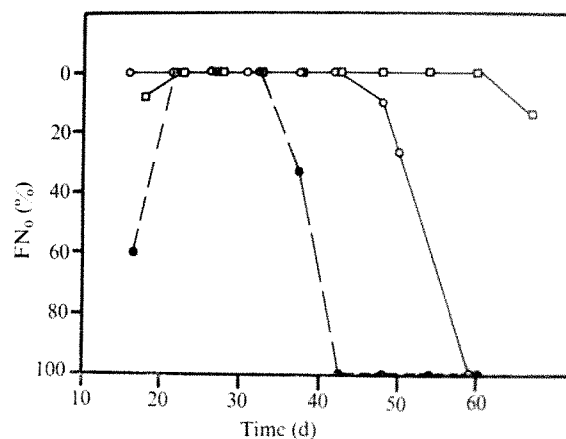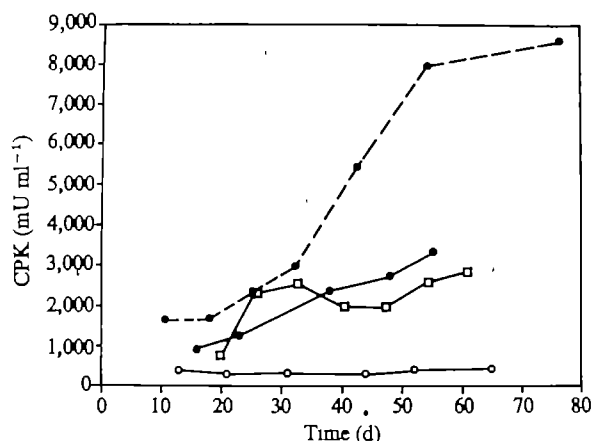
Fig. 5 Plasma CPK content of dystrophic chickens, untreated (broken line, mean of 18 birds per point), or treated with cyproheptadine (with FN — 5; ●, mean of 8 per point), or with both drugs (with FN = 5; □, mean of 9 per point). O, Normal chickens (mean of 14 per point).

apparent are illustrated in Table 1. Standard errors typical of the CPK determinations have been given elsewhere[3]. For all cases a highly significant difference was found between the untreated and the treated groups, with cyproheptadine over days 30–55, and with the combination of drugs from day 30 to the end of the study.

The two drugs used are already administered chronically in man for other conditions. In the mouse[8] LD$_{50}$ for toxicity of cyproheptadine given intraperitoneally is 55 mg kg^{-1}. In the chicken, it seems to be of still lower toxicity, perhaps denoting a more rapid removal or metabolism, since we have been unable to observe toxic symptoms at higher doses than that. We will report elsewhere more details of the conduct of the therapeutic trials, of histological evidence that parallels the other criteria, and of the lack of any effect in the dystrophic birds of these drug treatments on plasma enzymes unrelated to muscle disease.

There has been a long standing debate on the possibility of a change in blood circulation or of anoxia as a component in dystrophy (see refs 10 and 11). Our results may be attributable to antagonism of a 5-HT-induced effect on the microvasculature of dystrophic muscle, if such a component is indeed significant in inherited dystrophy; but that assumption is not as yet justified since the amelioration could, alternatively, arise by antagonism of an action of 5-HT elsewhere—for example, directly on the muscle[12]. In either case, the results suggest that an investigation in these chickens of a range of 5-HT antagonists could be of therapeutic interest, and they may provide a clue to one of the determinants of the dystrophic process.

M.S.H. thanks the Muscular Dystrophy Association of America for a Postdoctoral Fellowship. We thank Miss C. Pollina for assistance, and Drs D. W. Peterson and B. W. Wilson (Davis) for help with the supply of chickens: the flock used is maintained with the aid of MDAA.

Eric A. Barnard*
Arvind K. Bhargava†
Michael S. Hudecki

Department of Biochemistry,
State University of New York,
Buffalo, New York 14214

1 Asmundson, V. S., Kratzer, F. H., and Julian, L. M., *Ann. N. Y. Acad. Sci.*, 138, 49–58 (1966).
2 Julian, L. M , *Am J. Path.*, 70, 273–276 (1973).
3 Hudecki, M. S., and Barnard, E. A., *Res. Commun. Chem. Path. Pharmac.*, 14, 167–176 (1976).
4 Asmundson, V. S., and Julian, L. M., *J. Hered.*, 47, 248–252 (1956).

5 Farrell, P. M., Eyerman, E. L., and Turcon, L. L , *Ann. N. Y. Acad. Sci.*, 138, 102–111 (1966).
6 Holliday, T. A., Asmundson, V. S., and Julian, L. M., *Enzym. Biol Clin.*, 5, 209–216 (1965).
7 Dreyfus, J. C., Schapira, G., and Schapira, F., *Ann. N. Y. Acad. Sci*, 75, 235–249 (1958).
8 Boehringer–Mannheim, New York, catalogue no. 15926.
9 Woosher, M. L., and Optix, K , *Arch. Int Pharmacodyn*, 195, 252–259 (1972).
10 Bradley, W. G., et al., *Arch. Neur.*, 32, 466–473 (1975)
11 Mendell, J. R., Engel, W. K., and Derrer, E. C., *Science*, 172, 1143–1145 (1971).
12 Patton, B. M., Oliver, K. L., and Engel, W K., *Arch Neur.*, 31, 347–349 (1974)

Altered erythrocyte membrane phosphorylation in sickle cell disease

For several years evidence has been accumulating to indicate that the erythrocytes of individuals with sickle cell anaemia not only contain an abnormal haemoglobin but also exhibit several altered membrane characteristics. These include increased cation transport[1], possible alterations in phospholipid arrangement[2], and increased membrane rigidity[3]. The protein and glycoprotein content of membranes from erythrocytes of individuals with sickle cell anaemia does not seem to differ from that of normal erythrocyte membranes[4]. Many of the altered membrane characteristics appear to be a function of altered erythrocyte membrane structure due to membrane–haemoglobin interactions[5], however, whether or not other membranous protein–protein interactions are altered is not well documented.

We have recently reported[6] that human erythrocyte membranes contain, as well as a cyclic AMP-dependent protein kinase, another protein kinase which is capable of using either ATP or GTP as its phosphoryl donor. This enzyme is insensitive to regulation by cyclic nucleotides and catalyses the phosphorylation of membrane polypeptides in the area of bands 2 and 3 (nomenclature of Steck[7]). In this communication we report that the autophosphorylation of sickle-cell (SS) erythrocyte membranes by this enzyme differs from that observed with normal cells.

Blood was obtained from non-hospitalised individuals homozygous for haemoglobin S and from normal individuals through the Hematology Clinic at the University of Illinois

Fig. 1 Autophosphorylation of normal (a) and sickle cell (b) erythrocyte membranes in the presence of γ-[32P]P-GTP. Normal (38 µg) and sickle-cell (34 µg) erythrocyte membranes were incubated for 30 min at 37 °C in the presence of 0.1 M glycine-NaOH (pH 8.5), 10 mM MgCl$_2$, 0.2 mM γ-[32P]P-GTP (280 c.p.m. pmol^{-1}). The phosphopeptides were separated by electrophoresis in a 5% polyacrylamide slab gel (containing 0.2% SDS. The figure represents a densitometric tracing of an autoradiogram which was obtained after a 3-d exposure to the dried gel. Other experimental details were as previously described[6].
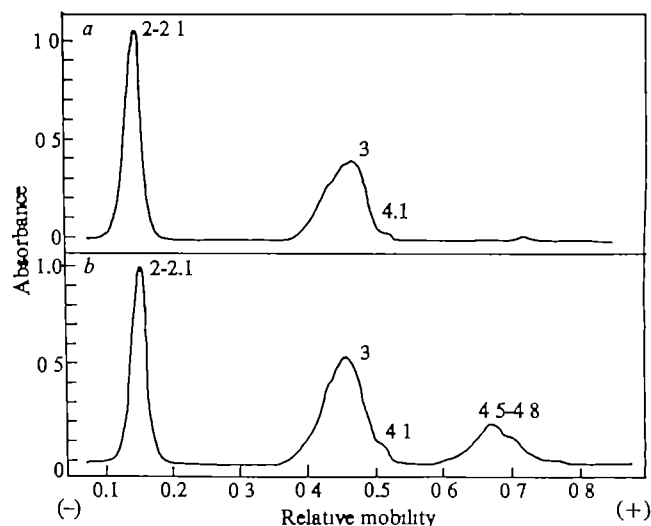
Phosphoryl donor		2–2.1	3	4.1	4.5–4.8
Table 1 Phosphorylation of normal and sickle-cell erythrocyte membrane polypeptides					
ATP	Normal (5)	1.01±0.05	1.43±0.11	0.05±0.01	0.12±0.02
	SS (5)	0.74±0.05*	1.51±0.13	0.12±0.03	0.24±0.02
ATP (+ cyclic AMP)	Normal (5)	1.27±0.11	1.61±0.13	0.16±0.02	0.43±0.05
	SS (4)	0.94±0.07	1.48±0.23	0.22±0.06	0.58±0.13
GTP	Normal (7)	0.64±0.05	0.57±0.13	0.01±0.01	0.01±0.01
	SS (8)	0.51±0.03†	0.83±0.12	0.07±0.04	0.20±0.03‡

Membranes were phosphorylated exactly as described in Fig. 1 except that γ-^{32}P-ATP (0.2 mM, 250 c.p.m. pmol[-1]) and cyclic AMP (2 μM) were present as indicated. When ATP was used as the phosphoryl donor, the incubation period was 5 min at 37 °C and the auto-radiogram was exposed for 4 d. The values were obtained by integrating the areas under the indicated peaks of densitometric tracings and are expressed as arbitrary densitometric units per μg membrane protein. All values are given as means ±s.e.m. The numbers in parentheses refer to the number of determinations. Statistical determinations were performed using the Student's t test.
* Significantly different from normal, $P < 0.01$.
† Significantly different from normal, $P < 0.05$.
‡ Significantly different from normal, $P < 0.001$.

Hospital. Haemoglobin-free erythrocyte membranes were prepared according to Dodge et al.[1] and stored in liquid nitrogen before use. Membranes were phosphorylated and applied to 5% polyacrylamide slab gels containing 0.2% SDS as previously described[2]. Autoradiograms which were prepared from the dried gels were scanned using a Zeineh Soft-Laser densitometer.

In confirmation of a previous study by others[4], we saw no large differences in the protein content of normal and SS erythrocyte membranes as analysed by SDS–polyacrylamide gel electrophoresis. When we studied the phosphorylation profile of the two types of membranes, however, we saw marked differences when γ-^{32}P-GTP was used as the phosphoryl donor (Fig. 1 and Table 1). As mentioned above, only one of the membrane-bound kinases can utilise GTP as its phosphoryl donor. As seen in Fig. 1 and Table 1, one or more polypeptides which migrate in the area designated as 4.5–4.8 are substrates for the GTP-utilising kinases in sickle-cell but not in normal erythrocyte membranes. This difference is not observed when ATP is used as the phosphoryl donor, either in the presence or absence of cyclic AMP (Table 1).

A second, albeit less marked, difference in sickle-cell membrane phosphorylation is that the phosphorylation of polypeptides 2–2.1 is slightly lower in sickle-cell membranes than in normal erythrocytes when either ATP (in the absence of cyclic AMP) or GTP is used as the phosphoryl donor. Although there is a tendency for the phosphorylation of 2–2.1 in the presence of ATP and cyclic AMP to be less in the SS than in the normal erythrocyte membranes, the data are not significantly different. In the presence of cyclic AMP no significant differences in the phosphorylation of normal and SS membranes are observed.

There are several possible explanations for the observed results. We favour the hypothesis that the conformation of kinase substrates may be altered in the SS membranes. It is conceivable that phosphoryl acceptor sites are covered (as for 2–2.1) or uncovered (as for 4.5–4.8) as a result of repeated sickling and unsickling during the lifespan of the erythrocytes. Alternately, the kinase itself may be altered such that its substrate specificity is decreased. Since the cyclic AMP-dependent protein kinases do not use GTP as a phosphoryl donor[6], it is unlikely that the phosphorylation of 4.5–4.8 in SS erythrocyte membranes in the presence of GTP is attributable to the cyclic AMP-dependent kinase. Further characterisation of the kinetics of the phosphorylation reactions in normal and SS membranes and the solubilisation of the kinases and their substrates is necessary to determine if one or more of these or other factors are responsible for the observed phenomenon.

It is well documented that several aspects of cation content and transport are altered in SS erythrocytes[1]. In addition, there are several lines of evidence implicating membrane phosphorylation in the regulation of cation

transport[10],[11]. Whether or not the altered membrane phosphorylation of SS erythrocytes plays a role in the altered transport processes in these cells or in other membrane phenomena remains to be determined.

We thank Dr E. Popescu for the sickle cells. This work was supported by the American Cancer Society. M. Hosey is the recipient of a National Institutes of Health Fellowship. M. Tao is an Established Investigator of the American Heart Association.

M. MARLENE HOSEY
MARIANO TAO

University of Illinois at the Medical Center,
Department of Biological Chemistry,
835 South Wolcott Avenue,
Chicago, Illinois 60612

1 Kuranstin-Mills, J., Kudo, M., and Addae, S K., Clin. Sci. molec. Med., 46, 679–692 (1974).
2 Gordesky, S. E., Marinetti, G. V., and Segal, G. B., Biochem. biophys. Res. Commun., 47, 1004–1009 (1972).
3 Chien, S., Usami, S., and Bertles, J. F., J. clin. Invest., 49, 623–634 (1970).
4 Durocher, J R., and Conrad, M. E., Proc. Soc. exp. Biol. Med, 146, 373–375 (1974).
5 Lorand, L. S., Wallas, C. H., and Weems, H., in Proc. First National Symposium on Sickle Cell Disease (edit. by Hercules, J. I., Schechter, A. N., Eaton, W. A., and Jackson, R. A.), 213–214 (DHEW Publication No (NIH) 75-723, Bethesda, 1974).
6 Hosey, M. M., and Tao, M., Biochemistry, 15, 1551–1559 (1976).
7 Steck, T. L., J. molec. Biol., 66, 295–305 (1972).
8 Dodge, J. T., Mitchell, C., and Hanahan, D. J., Archs Biochem. Biophys., 100, 119–130 (1963).
9 Tao, M., and Hackett, P., J. biol. Chem., 248, 5324–5332 (1973).
10 Rudolph, S. A., and Greengard, P., J. biol. Chem., 249, 5684–5687 (1974).
11 Cha, Y.-N., and Lee, K.-S., J. gen. Physiol., 67, 251–261 (1976).

Inhibition of intercellular adhesion in a cellular slime mould by univalent antibody against a cell-surface lectin

THE cellular slime moulds, *Dictyostelium discoideum* and *Polysphondylium pallidum* contain carbohydrate-binding proteins (lectins) that increase significantly in amount in soluble crude extracts as amoebae differentiate from the vegetative stage (when the cells do not associate) to the aggregation-competent stage (when the cells can form stable intercellular contacts)[1-3]. Lectins from the two species, assayed as agglutinins of erythrocytes, have been purified by affinity techniques and, on the basis of their carbohydrate-binding specificities, as well as several physicochemical properties, have been shown to be distinct proteins[3-5]. Distinct lectin activities have also been identified in four other species of slime moulds[6]. The lectins from *D. discoideum* and *P. pallidum* accumulate on the cell surface with development of aggregation competence as demonstrated by erythrocyte-rosette experiments[1,2,7], immunofluorescent and immunoferritin techniques[7,8], and lactoperoxidase-catalysed iodination of intact cells[8]. Both species also contain cell-
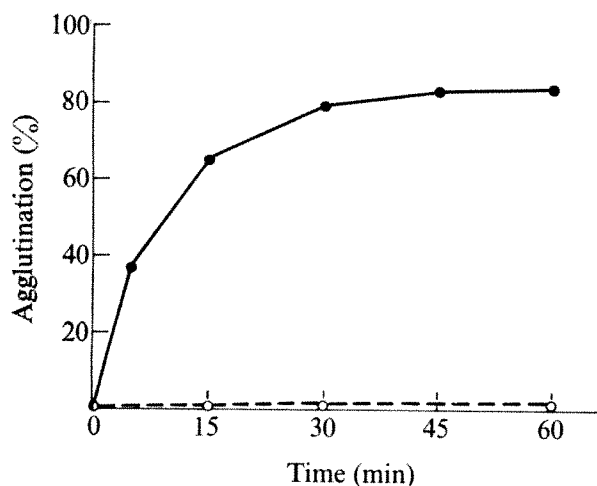
Fig. 1 Kinetics of agglutination of aggregation-competent ameobae (●) and vegetative amoebae (○). Vegetative amoebae and aggregation-competent amoebae were collected after 55 h and 96 h of incubation on agar, respectively. Cells were washed several times in cold water to separate them from bacteria and were then suspended at 9×10^6 ml^{-1} in EDTA–phosphate buffer (16.7 mM Na$_2$HPO$_4$–KH$_2$PO$_4$, 10 mM EDTA, pH 6.2) containing 0.2 M D-glucose and 1 mg ml^{-1} bovine serum albumin (BSA). The cell suspension was dispersed into single cells by repeated Manostat pipetting. After 30 min in an ice bath, the cell suspension was again dispersed as above. A 500-μl sample of the cell suspension was added to each of several individually cut out wells of a Linbro plate (16 mm diameter, FB–54). The wells were gyrated at 115 r.p.m. on a New Brunswick G-24 shaker (radius of gyration 3/4 inch) at 22 °C. At various times, the contents of wells were carefully poured into plastic vials containing 20 ml of EDTA–phosphate buffer. The vials were gently mixed and after 5 min, to allow osmotic re-equilibration, single cells were counted with a Coulter Counter (Model ZBI, 100-μm aperture, 1/amplitude = 1, 1/aperture current = 0.354, impedance = 10 K, threshold = 10–60). With these settings, approximately 90% of the single cells were counted. The total number of cells present in the suspension was determined by dispersing the agglutinates into single cells by Manostat pipetting and counting at the above settings. The percentage agglutination was computed as the percentage of single cells that entered into aggregates. Two wells were used for each time point. Readings on independent samples agreed to within 5%.

surface receptors with a very high affinity for the homotypic lectin and a lesser affinity for the heterotypic lectin[10], suggesting a possible basis for species-specific intercellular affinities. In the case of *D. discoideum*, we have shown that the high affinity receptor appears on the cell surface as amoebae differentiate into the cohesive state[10].

This series of findings raises the possibility that intercellular adhesion in the cellular slime moulds is mediated by the association of cell-surface lectins with complementary surface receptors, presumed to contain complex oligosaccharides. This possibility is supported by the finding that cohesiveness of aggregation-competent *P. pallidum* cells is inhibited by specific simple sugars that react with pallidin, the lectin derived from this species[2]. High concentrations of these sugars are, however, required (0.1 M or higher), suggesting the possibility of nonspecific effects on unknown components of the cell, rather than inhibition of cell-surface pallidin. To test the possibility that pallidin on the surface of *P. pallidum* amoebae is indeed involved in the developmentally regulated cohesiveness of these cells, we raised antiserum to purified pallidin and determined the effect of Fab fragments, derived from the antibodies, on the cohesiveness of *P. pallidum* cells. We now report that these univalent antibodies inhibit the cohesiveness of differentiated amoebae.

Polysphondylium pallidum strain WS 320 was grown in the dark in the presence of *Aerobacter aerogenes* on

nutrient agar at 22 °C with 10^5 spores inoculated on each 100-mm agar plate, as described previously[2]. In these conditions, the amoebae became aggregation competent by 96 h when the bacteria had been substantially cleared. The absence of light prevents the cells from aggregating and proceeding through culmination. On illumination, however, aggregation of the cells is apparent within 30 min. Vegetative cells were obtained after 55 h of culturing, when a thick carpet of bacteria still covered the agar. These cells, once free of bacteria, require a several-hour interphase period before they are aggregation competent. Cells of both stages were collected and washed as previously described[2]. An index of cell cohesiveness was obtained by gyrating a suspension of mechanically dispersed cells and determining the percentage reduction in single cells as they agglutinated. In the hypertonic assay conditions used (see Fig. 1), only the aggregation-competent cells formed agglutinates, whereas vegetative amoebae remained as single cells. The critical factors in this assay, including a morphological study of the agglutinates by transmission and scanning electron microscopy, will be presented later.

Pallidin was purified by affinity adsorption[5]. The purified protein gave a single band on polyacrylamide electrophoresis in the presence of sodium dodecyl sulphate and 2-mercaptoethanol, even when as much as 200 μg of protein was applied to the gel. Antiserum, raised in a rabbit using a conventional procedure as described previously[7], formed a single precipitin band on immunodiffusion against either
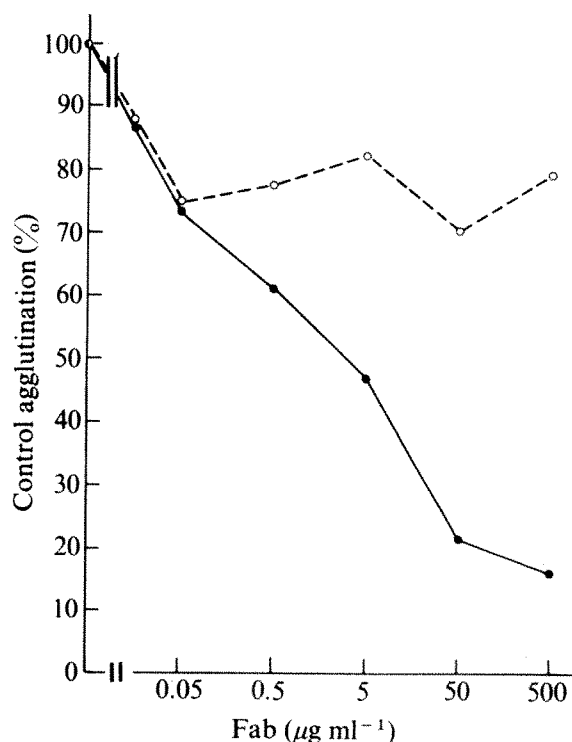
Fig. 2 Effects of immune (●) and normal Fab (○) on *P. pallidum* agglutination. Cells were collected after 96 h of incubation, washed and suspended to 1.8×10^7 ml^{-1} in EDTA–phosphate buffer containing 0.4 M D-glucose and 2 mg ml^{-1} BSA. Aliquots of the cells were dispersed and mixed with equal volumes of varying concentrations of normal or immune Fab diluted in EDTA–phosphate buffer. After a preincubation period of 30 min at 0 °C, the cells were dispersed again and added to Linbro plastic wells, 500 μl per well. After 30 min of gyration by which time equilibrium was approached, the percentage of cells that entered into agglutinates was determined as above (see Fig. 1 legend). For each of a series of concentrations of normal and immune Fab, the percentage of the control level of agglutination in the absence of Fab was computed. Two replicate wells were used per treatment. Protein concentration was determined by the Lowry method using BSA as a standard.

purified pallidin or crude extracts of *P. pallidum* cells. IgG was purified by ammonium sulphate precipitation followed by DEAE–cellulose chromatography[11]. Fab fragments were derived by papain digestion and purified by the standard procedure[12]. Control Fab fragments were prepared in parallel from normal rabbit serum.

The Fab fractions were assayed for their ability to inhibit haemagglutination produced by purified pallidin in a standard assay. The reaction mixture added to the wells of a Microtier V plate was as follows: 25 μl of a twofold dilution series of pallidin in 75 mM NaCl, 75 mM Na_2HPO_4–KH_2PO_4, (pH 7.2 PBS), 25 μl of formalinised human O erythrocytes[2] 2.5% (v/v) in PBS; and 25 μl of Fab diluted in 16.7 mM Na_2HPO_4–KH_2PO_4, 10 mM EDTA (pH 6.2). At a final concentration of 0.5 μg ml^{-1}, immune Fab reduced the haemagglutination titre by 50%: normal Fab did not produce inhibition at concentrations up to 250 μg ml^{-1}.

Immune Fab (at 3 μg ml^{-1}) directed against pallidin inhibited the cohesiveness of *P. pallidum* cells by 50%, whereas control Fab had little effect (Fig. 2). Inhibition of agglutination was also obvious on microscopic inspection.

Beug and colleagues[13,14] have demonstrated that Fab fragments from antisera against crude fractions of differentiated *D. discoideum* amoebae block cohesion of *D. discoideum* cells; but in this case the cell-surface antigens with which the contact-blocking Fab fragments react are not known. In the present experiments, Fab fragments directed against the cell-surface lectin, pallidin, were shown to inhibit cohesiveness of differentiated *P. pallidum* cells, as measured in a gyrated suspension. Although further experiments are required to fully define the morphogenetic role of the cellular interactions measured by our cohesiveness assay, we assume that this experimentally defined cohesiveness is functionally significant, since it is correlated with the ability of cells at different developmental steps to form stable intercellular contacts when moving on a solid substratum. Together with other evidence summarised here and reviewed elsewhere[15], the present experiments provide strong support that stage-specific intercellular adhesion is mediated by the cell-surface lectin, pallidin.

STEVEN D. ROSEN*
PATRICIA L. HAYWOOD
SAMUEL H. BARONDES

Department of Psychiatry,
School of Medicine,
University of California, San Diego,
La Jolla, California 92093

1 Rosen, S. D., Kafka, J. A., Simpson, D. L., and Barondes, S. H., *Proc. natn. Acad. Sci. U.S.A.*, 70, 2554 (1973).
2 Rosen, S. D., Simpson, D. L., Rose, J. E., and Barondes, S. H., *Nature*, 252, 128 (1974).
3 Frazier, W. A., Rosen, S. D., Reitherman, R. W., and Barondes, S. H., *J. biol. Chem.*, 250, 7714 (1975).
4 Simpson, D. L., Rosen, S. D., and Barondes, S. H., *Biochemistry*, 13, 3487 (1974).
5 Simpson, D. L., Rosen, S. D., and Barondes, S. H., *Biochim. biophys. Acta*, 412, 109 (1975).
6 Rosen, S. D., Reitherman, R. W., and Barondes, S. H., *Expl Cell Res.*, 95, 159 (1975).
7 Chang, C-M, Rosen, S. D., and Barondes, S. H., *Expl Cell Res.* (in the press).
8 Chang, C-M, Reitherman, R. W., Rosen, S. D., and Barondes, S. H., *Expl Cell Res.*, 95, 136 (1975).
9 Siu, C-H, Lerner, R. A., Ma, G., Firtel, R. A., and Loomis, W. F., *J. molec. Biol.*, 100, 157 (1976).
10 Reitherman, R. W., Rosen, S. D., Frazier, W. A., and Barondes, S. H., *Proc. natn. Acad. Sci. U.S.A.*, 72, 3541 (1975).
11 Campbell, D. H., Garvey, J. S., Cremer, N. E., and Sussdorf, D. H., *Methods in Immunology* (Benjamin, New York, 1970).
12 Porter, R. R., *Biochem. J.*, 73, 119 (1959).
13 Beug, H., Gerisch, G., Kampff, S., Riedel, V., and Cremer, G., *Expl Cell Res.*, 63, 147 (1970).
14 Beug, H., Katz, F. E., and Gerisch, G., *J. Cell Biol.*, 56, 647 (1973).
15 Barondes, S. H., and Rosen, S. D., in *Neuronal Recognition* (edit. by Barondes, S. H.), (Plenum, New York, 1976).

New approach to determination of specific functions of platelet membrane sites

IT has been suggested that the layer of bound carbohydrate on the surface of a platelet contains groupings vital to the haemostatic function of the platelet[1]. This haemostatic role is largely dependent on the ability of the platelet to adhere to exposed subendothelial components in the event of vessel damage[2]. Platelet aggregation, induced by bovine factor VIII and ristocetin, and platelet adhesion to subendothelium are impaired in the Bernard–Soulier syndrome, while aggregation induced by ADP is apparently normal[3,4]. A gross abnormality in the 155,000-molecular weight glycoprotein seen in membrane fractions prepared from platelets from this syndrome[5] suggests that the glycoprotein acts as an acceptor–receptor during the interation with subendothelium and macromolecular aggregation-inducing agents. Here we report further evidence that a specific site on the platelet is required for ristocetin and bovine factor VIII to induce platelet aggregation and for platelet adhesion when aggregation and platelet adhesion to subendothelium are inhibited by an acquired antiplatelet antibody found in a polytransfused Bernard–Soulier patient. This antibody induced *in vitro*, on normal human platelets, a specific Bernard–Soulier-like defect, so that adhesion to subendothelium was impaired and aggregation with ristocetin and bovine factor VIII was defective, while ADP-mediated aggregation was unaffected.

The antibody, which was shown to be an IgG, was isolated from the plasma of a Bernard–Soulier patient (P) who had received multiple transfusions of platelet concentrates during the arrest of a bleeding episode. It seemed to differ from the usual antiplatelet alloantibodies: a negative complement fixation test was observed with a panel of 25 control human

Fig. 1 Effect of different dilutions of the antibody from patient P on the aggregation of control human PRP induced by ADP, collagen, bovine factor VIII and ristocetin. Citrated control human PRP (0.2 ml) was incubated for 60 s in the cuvette of a Born MK II miniaggregometer (produced by the Department of Pharmacology, Royal College of Surgeons, London) in the presence of several different non-agglutinating dilutions of the antibody (in saline) before addition of 10 μl of one of the following inducers (final concentration): ▲, 0.6 μmol of ADP (Sigma); ×, 20 μg of collagen (Stago, Asnières); □, 160 μg of bovine factor VIII (AB Kabi, Stockholm); ○, 200 μg of ristocetin (Lundbeck, Copenhagen). Results (mean values) of 10 experiments are expressed as the percentage inhibition of the aggregation intensity in the absence of antibody.
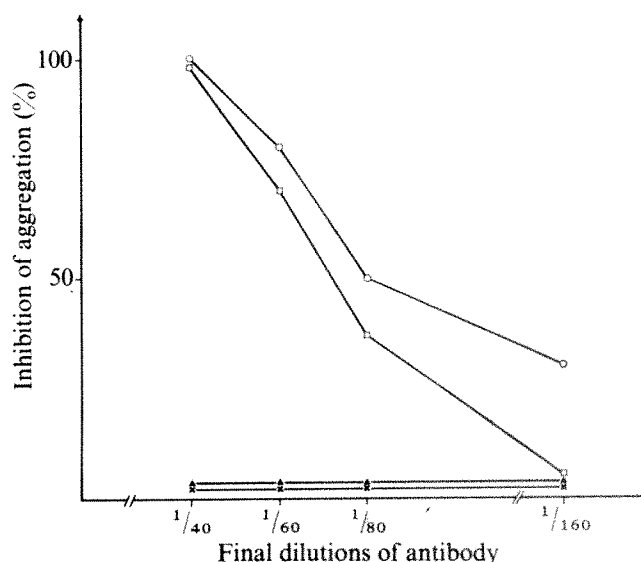

Final dilutions of antibody

Table 1 Consumption studies of the antibody from patient by platelet stroma

Antibody final dilution	Residual agglutinating activity with control PRP after incubation with				Residual inhibitory effect on ristocetin-induced aggregation of control PRP after incubation with			
	Saline	Control platelet stroma	B S 1 platelet stroma	B S 2 platelet stroma	Saline	Control platelet stroma	B S 1 platelet stroma	B S 2 platelet stroma
1:20	50	12	60	55				
1:40	15	0	18	16	80	10	45	50
1:80					45	5	20	45

Hypercitrated PRP from control human or two other Bernard-Soulier patients, BS1 and BS2, were centrifuged for 10 min at 2,000g and 15 °C. Platelets pellets were resuspended in 0.015 M Tris-HCl buffer, pH 7.4, containing 0.135 M NaCl, 0.053 M KCl and 0.015 M EDTA. These pellets washed twice in this buffer, then finally resuspended in the same buffer without EDTA, were disrupted by five times freeze–thawing and washed twice in saline (27,500g, 15 min, 15 °C) to remove most of the intracellular constituents. The stromas were resuspended at a concentration of 1 mg of protein per ml in saline. Samples (0.2 ml) of this membrane-enriched fraction or 0.2 ml of saline were incubated with 0.2 ml of different dilutions of 6 mg of the antibody solution (final dilutions of the antibody were 1:20, 1:40, 1:80) for 120 min at 37 °C, followed by overnight incubation at 4 °C. After centrifugation at 27,500g for 15 min at 15 °C, the supernatants were tested for their residual agglutinating activity (left hand side) on control PRP (results expressed as percentage of agglutination) and inhibitory effect on ristocetin-induced aggregation (right hand side) of control PRP (results expressed as percentage of inhibition of the intensity of ristocetin aggregation). Mean values of three experiments.

platelets by a microtechnique described before[6], the lymphocytotoxicity test of Terasaki and McClelland[7] revealed only a weak anti-HLA-A1 antibody, and the thromboagglutination test[8] was positive on 20 control human platelets whatever the HLA, PLA$_1$, KOa or KOb type. When control citrated platelet-rich plasma (PRP) from 50 donors was incubated at 37 °C and stirred in a Born MK II miniaggregometer cuvette in the presence of the antibody, agglutination was observed after about 1 min. The intensity of the agglutination of control PRP was directly related to the concentration of antibody added. No agglutination was observed with a final dilution greater than 1:40. PRP from a patient with severe quantitative von Willebrand disease was normally agglutinated by the antibody. In contrast no agglutination effect (at any dilution) was observed when the antibody was added to PRP obtained from two other Bernard–Soulier patients. Using PRP from the family of one of these patients, we found that the agglutinating effect of the antibody from patient P was nil for an apparently healthy brother, about 60% of normal for another healthy brother and also for the mother and the father, making possible a distinction between the normal and heterozygous state.

Figure 1 shows that the antibody from patient P was a strong inhibitor of aggregation induced by ristocetin and bovine factor VIII. The inhibitory effect was related to the concentration of antibody used. In contrast, ADP or collagen-induced aggregation were not inhibited by the antibody. Table 1 shows that after incubation with control platelet stroma, both agglutinating activity and inhibition of ristocetin-induced aggregation by the antibody were suppressed, while with the platelet stroma of the two other Bernard–Soulier patients,

Table 2 Effect of antibody from patient P on interaction of platelets with rabbit aorta subendothelium

Antibody final dilution	Interaction of platelets with aorta subendothelium	
	Naked	Adhesion
0	28±12	72±12
1:80	43	57
1:40	76	24

The procedure was as described by Baumgartner[2]. Everted segments of rabbit aorta, previously denuded of endothelium were mounted on the central rod of a perfusion chamber. Citrated whole blood from donors was first incubated for 10 min with the antibody (final dilutions 1:40, 1:80) or with saline before being circulated for 10 min at 37 °C with an average flow rate of 160 ml min^{-1} through the chamber. After the chamber had been rinsed with phosphate buffer, the blood vessel segments were fixed and embedded in Epon for examination by light microscopy and for evaluation of platelet interaction with subendothelium. Results are expressed as the percentage of the subendothelial surface that is devoid of platelets (naked) or that is covered by adherent platelets (adhesion). Results of control without antibody are given as the mean values of 10 experiments±1 s.d.

much of the antibody activity was recovered in the supernatant. In the presence of antibody dilutions that inhibited 100% of control PRP ristocetin-induced platelet aggregation, platelet adhesion to subendothelium was estimated by the Baumgartner technique[2] to be 24% compared with 72% without antibody (Table 2).

The agglutinating effect of the antibody from patient P on control human PRP and not on the two other Bernard–Soulier PRP, could indicate that the antibody was an antiplatelet antibody directed against an antigen present on all normal platelets but absent, very reduced or abnormal on Bernard–Soulier platelets. This suggestion is supported by the fact that control human platelet stromas consumed the antibody while Bernard–Soulier platelets were much less efficient in doing so. Using subagglutinating dilutions of the antibody from patient P, we have demonstrated the strong and specific inhibition of aggregation induced by ristocetin and bovine factor VIII while ADP and collagen aggregation were not influenced. In the presence of the antibody, adhesion of normal platelets to subendothelium was much reduced; this may indicate that the antigen recognised by the antibody was necessary for the aggregation inducing agents bovine factor VIII and ristocetin, and for adhesion to subendothelium, but not for aggregation mediated by ADP and collagen. The antibody was not directed against von Willebrand factor, for it induced normal agglutination of PRP from severe quantitative von Willebrand disease. Furthermore, von Willebrand factor activity in ristocetin-induced platelet aggregation was normal in the fraction eluted in the void volume after gel filtration on Sepharose 2B of plasma from patient P[3].

As the platelets of patient P had a surface defect involving a membrane glycoprotein[4], we suggest that the antibody is directed against a specific platelet site, implicated in the interaction with subendothelium and macromolecular aggregation-inducing agents. This antibody makes possible a distinction between the surface groupings involved in aggregation induced by ADP and collagen on the one hand, and in aggregation induced by ristocetin and bovine factor VIII and adhesion to subendothelium on the other hand. The platelet antigen recognised by the antibody could be the 155,000 molecular weight glycoprotein, but the exact antigenic site of the molecule remains unknown and is under investigation.

G. TOBELEM
S. LEVY-TOLEDANO
R. BREDOUX
H. MICHEL
A. NURDEN
J. P. CAEN

INSERM, U150,
CNRS, ERA 335,
Hôpital Lariboisière,
2, rue Ambroise Paré,
75475 Paris Cedex 10, France

L. DBGOS*

INSERM, U93,
Hôpital Saint-Louis,
2, place du Dr Alfred Fournier,
75475 Paris Cedex 10, France

1 Nurden, A. T., and Caen, J. P., *Thromb. Haemostasis*, 35, 139–150 (1976).
2 Baumgartner, H. R., *Thromb. Diath. Haemorrh.*, Suppl. 11, 161–176 (1972).
3 Caen, J. P., *et al., J. Lab. clin. Med.*, 87, 586–596 (1976).
4 Weiss, H. J., Tschopp, T. B., Baumgartner, H. R., Sussman, I. L., Johnson, M. M., and Egan, J. J., *Am. J. Med.*, 57, 920–931 (1974).
5 Nurden, A. T., and Caen, J. P., *Nature*, 255, 720–722 (1975).
6 Colombani, J., D'Amaro, J., Gabb, B W., Smith, G., and Sbejgoard, A., *Transplant. Proc.*, 3, 121–126 (1971).
7 Turaski, P. I., and McClelland, J D., *Nature*, 204, 998–1000 (1964).
8 Daussot, J., and Colin, M., *Rev. fr. Etud. clin. biol.*, 3, 60–68 (1958)

Kinetics of agonist-induced intrinsic fluorescence changes in membrane-bound acetylcholine receptor

ACETYLCHOLINE receptor(AChR)-rich membranes derived from fish (*Electrophorus* and *Torpedo*) electric organs have been used extensively for *in vitro* binding and functional studies[1]. These membrane fragments (microsacs) are particularly suited for physical investigations of AChR–ligand interactions because of their inherently high content of receptor, retained in its natural environment. A number of extrinsic fluorescent probes have been applied to the determination of binding specificity, equilibria and, in a preliminary way, kinetics at the level of the isolated membrane, as well as the neuromuscular junction and purified receptor[1–4]. We report here the existence of intrinsic protein fluorescence changes, induced in membrane fragments from *Torpedo marmorata* electroplaques by the natural neurotransmitter acetylcholine. We have used highly sensitive fluorescence kinetic techniques to obtain the concentration dependence and some of the corresponding equilibrium and
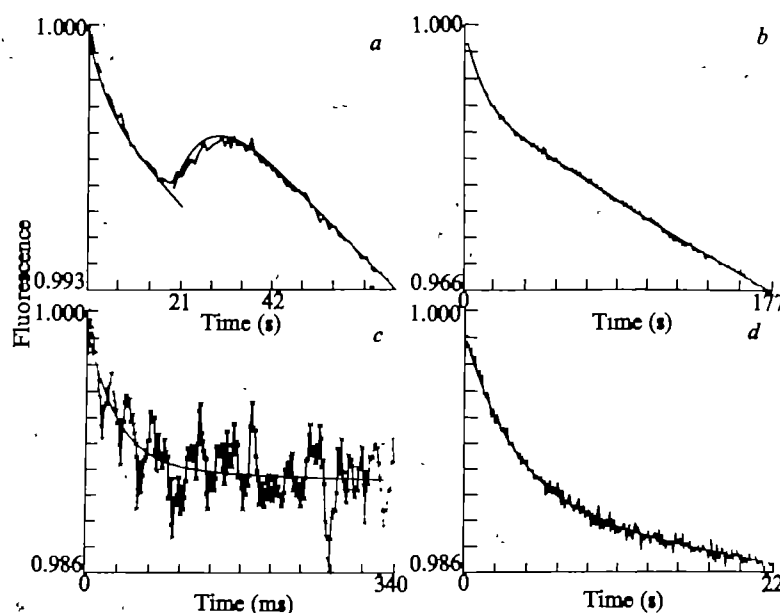
rate constants for the initial binding and subsequent isomerisation reactions. The latter can be correlated with electrophysiological data and the functional properties of the post-synaptic membrane[7–9], particularly with respect to the process of desensitisation, and with other quantitative results obtained with the same microsac preparations[10]. In particular, our data provide physical evidence for the existence of at least three conformational states of the cholinergic receptor protein *in situ*, interrelated by reversible kinetic processes.

AChR-rich microsacs were prepared essentially as described by Cohen *et al.*[11], and their α-toxin-binding capacity determined by the Millipore filter assay[1a]. Fluorescence kinetic experiments were carried out in the stopped-flow apparatus designed in this laboratory[1a]. Further experimental details are given in the legends to Figs 1 and 2.

Figure 1 shows the changes in the intrinsic fluorescence of the membrane-bound AChR induced by mixing with the natural neurotransmitter, ACh, and another cholinergic agonist, carbamylcholine. An exponential decay in fluorescence was observed on the addition of more than 10 μM ACh to the microsacs (Fig. 1a). This initial response reversed to the original fluorescence decay after a delay which increased with increasing ACh concentrations or with a decreasing acetylcholinesterase (AChE) content of the microsacs. The decay, and its reversal, were superimposed on a uniform linear decrease in fluorescence emission probably due to photolysis caused by the high illumination intensities in the stopped-flow apparatus. (The drift could be determined easily and corresponded to a -0.012% s^{-1} relative change in fluorescence—that is, much slower than the ligand-induced effect. All records were corrected appropriately.)

At ACh concentrations < 2 μM, the initial response was no longer monoexponential and its magnitude was markedly reduced, presumably due to rapid degradation of the ligand by the AChE present in the partially purified microsacs. In the presence of small amounts of the AChE inhibitor paraoxon

(diethyl-*p*-nitrophenylphosphate, 150 nM), however, the ACh-induced fluorescence quenching could be followed over a wide concentration range (Figs 1*b* and 2).

The relaxation times and amplitudes characterising the progress curves after mixing were concentration dependent but reached plateau values at high ACh levels (Fig. 2). Over a limited concentration range, an additional faster exponential decay in fluorescence was seen (Figs 1*c* and 2) but was not detectable elsewhere, for reasons given below.

Similar kinetic phenomena, except for the reversal were observed with another agonist, carbamylcholine, which is not hydrolysed by the esterase (Fig. 1*d*). The relaxation time and amplitude, however, reached constant values at significantly higher concentrations of carbamylcholine as compared with ACh (65 against 4 μM, respectively).

The specificity of the ligand-induced fluorescence changes and their attribution to distinct states of the receptor protein follow from findings that: the effects were abolished by pre-incubation with a tenfold molar excess of α-cobrotoxin; and the total amplitude of the fluorescence quenching was proportional to the specific activity (in terms of toxin binding) of various microsac preparations and not to the total protein or AChE content. The essential reversibility of the phenomenon—that is, the return of the AChR to its initial resting state—is attested to by the reversal observed as a consequence of dissociation and enzymatic hydrolysis of the free transmitter. The onset time and time course of the reversal were in good agreement with expectation on the basis of AChE determinations using acetylthiocholine[14].

The simplest reaction mechanism consistent with our experimental results consists of a fast binding step followed by a slow isomerisation

$$A + R \underset{k_2}{\overset{k_1}{\rightleftharpoons}} AR \underset{k_4}{\overset{k_3}{\rightleftharpoons}} AD \qquad (1)$$
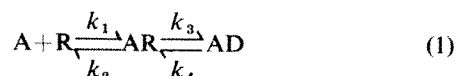
where A is the agonist and R and D are two distinct conformers of the AChR. The experimentally observed slow relaxation time, τ_{II}, corresponds in this sequential scheme to the equilibration of the second step to which the first is thermodynamically coupled

$$(\tau_{II})^{-1} = \left[\frac{A}{K_1 + A}\right] k_3 + k_4 \qquad (2)$$

where K_1 is the dissociation constant (k_2/k_1). The reciprocal relaxation (Fig. 2) has an ordinate intercept of k_4, an initial slope of k_3/K_1 and a limiting value at high [A] of $k_3 + k_4$. The corresponding amplitude (in terms of relative fluorescence) is given by

$$\alpha_{II} = \alpha_{total} - \alpha_I = \left[\frac{A}{K_1 + A(1 + K_2)}\right](\Delta\Phi_1 + K_2\Delta\Phi_2)$$
$$- \left[\frac{A}{K_1 + A}\right]\Delta\Phi_1 \qquad (3)$$

where $\Delta\Phi_1$ and $\Delta\Phi_2$ are the fractional changes in fluorescence on passing from state R to states AR and AD, respectively, and K_2 is the isomerisation equilibrium constant (k_3/k_4). α_{II} has an ordinate intercept of zero, an initial slope of $K_2\Delta\Phi_2/K_1$, a maximum magnitude at $[A] \sim K_1(K_2\Delta\Phi_1/\Delta\Phi_2)^{-\frac{1}{2}}$ and a plateau value of $K_2(\Delta\Phi_2 - \Delta\Phi_1)/(1 + K_2)$.

Using the above formulation, it was possible to fit the data obtained at various concentrations of ACh or carbamylcholine with one set of parameters for each agonist (Fig. 2 and Table 1). It is seen that carbamylcholine has a lower affinity for the R state of the receptor but the transition from the AR to the AD state is more rapid and extensive (k_3 and K_2 larger than for

ACh). The fluorescence parameters, however, are about the same for both drugs, and indicate that the emission of the receptor in the states R, AR, and AD decreases in that order. The values of K_1 for both ACh and carbamylcholine agree well with the apparent dissociation constants derived from [²²Na] efflux measurements in the microsac preparation and from direct binding to solubilised receptor[10].

The limited data for the rapid decay process (Figs 1 and 2) do not suffice for the determination of the corresponding rate constants, although k_1 is certainly $> 10^6$ M^{-1} s^{-1}. Thus, at low ACh concentration (for example, 1 μM), α_I is too small ($\leqslant 0.2\%$), whereas at high concentrations (> 40 μM), τ_I is too small (< 10 ms) for the relaxation to be perceptible directly. Its presence is indicated nevertheless by the decrease in α_{II} at high levels of ACh (Fig. 2).

The addition of a further rapid isomerisation step between an inactive complex AR (equation 1) and a conducting form AR′ does not alter the ability of the above formulation to account for the data of Fig. 2 (except to change the quantitative interpretation of K_1 and k_3). Such a scheme is required to rationalise the submillisecond rise time of the endplate potential[8,15] and the analyses made of potential fluctuations (membrane noise)[16,17]. If the AR–AR′ equilibrium is rapidly established and particularly if it favours the AR state[17], we would fail to detect it due to the 3–5 ms time resolution of the conventional stopped-flow apparatus[31], even if a distinct fluorescence change were involved.

The second relaxation process is most likely related to the formation of an inactive or 'desensitised' state of the receptor (AD), a phenomenon which occurs in *T. marmorata in vivo*[18] and *in vitro*[10,19,20] and, moreover, generally in the cholinergic synapse[8,9,21–24]. The rate constant k_3 (Table 1) indicates that desensitisation occurs in the time range of seconds, a finding compatible with most electrophysiological data. In addition, other studies (unpublished results) have shown the expected influence of temperature and of effectors known to influence desensitisation on the fluorescence relaxation process. Although the simple scheme outlined above accounts well for the experimental results, the rapid reversal of the fluorescence decay on ACh hydrolysis (Fig. 1*a*) suggests the existence of an alternative route for recovery from the desensitised state with a rate > 0.1 s^{-1} (and thus incompatible with the value of k_4, Table 1). This phenomenon, which may to a limited degree represent the physiological situation, is consistent with a cyclic scheme

Fig. 2 Dependence on ACh concentration of the fluorescence decay rates (τ^{-1}) and amplitudes (α) after mixing AChR-rich microsacs with ACh in the stopped-flow fluorimeter. Specific activity: 1,000 pmol α-toxin-binding sites per mg protein. AChR concentration after mixing was 20 nM in α-toxin sites. ■, Decay rate measured in the presence of the AChE inhibitor paraoxon (300 nM); □, decay rate measured in the absence of the AChE inhibitor; ●, observed amplitude of decay in fluorescence in the presence of 300 nM paraoxon; ○, amplitude measured in the absence of paraoxon; ▲, amplitude of fast decay in fluorescence observed at 4 μM ($\tau_I = 120$ ms) and at 12 μM ($\tau_I = 32$ ms) ACh, with the same membrane preparation used in experiment shown in Fig. 1*c*. Solid lines correspond to the theoretical curves calculated for the sequential scheme (equations (1) and (2)) using the values given in Table 1.
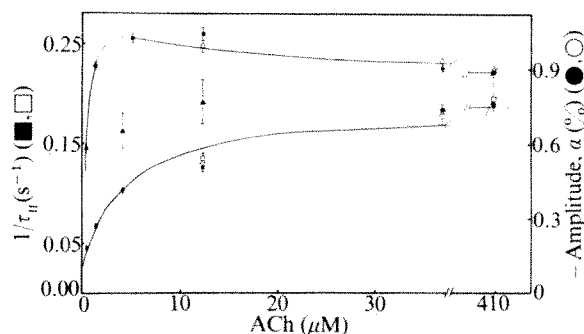
Table 1 Parameters derived from the ligand-induced fluorescence changes using the sequential scheme outlined in equations 1–3

	ACh*	Carbamylcholine
K_1	4.1 µM	66 µM
	2.0 µM†	50 µM†
k_3	0.16 s^{-1}	0.37 s^{-1}
k_4	0.027 s^{-1}	0.01 s^{-1}
K_2	6	37
$\Delta\Phi_1$	−0.67%‡	−0.75%
$\Delta\Phi_2$	−1.7%	−1.5%

*In the presence of 150 nM AChE inhibitor, paraoxon. Temperature: 20 °C.

†From Table 2 in ref. 10.

‡This and related values have been normalised to a specific activity of 1 nmol ^{3}H-α-toxin per mg protein. The measured fractional fluorescence changes, $\Delta\Phi_1$ and $\Delta\Phi_2$, are related to the actual relative changes in the fluorescence of the AChR by a proportionality factor β which denotes the fractional contribution made by the receptor to the total protein emission. That is, the true changes associated with the intrinsic AChR fluorescence are given by $\beta\Delta\Phi_1$ and $\beta\Delta\Phi_2$, where β could have a value ≥ 10.

extending equation (1) to include the dissociation of the AD complex and the isomeric transition of the D to the R state. Such a mechanism, which was considered originally by Katz and Thesleff[21] and developed further by Rang and Ritter[22,23] best fits available pharmacological data[8,9]. Our results indicate that the desensitised receptor has a much higher affinity for the agonist than does the resting state ($K_3 < K_1/K_2 \ll K_1$, where K_3, the dissociation constant for the AD state, must be < 0.3 µM to account for the consistency of the data with the sequential scheme).

It is important to note that regardless of the true binding order associated with activation (that is, number of ligand molecules required per functional receptor, a value which may well exceed one[7–9,25–27]) the isomerisation step (desensitisation) we have observed seems to proceed from a mono-liganded state (we cannot exclude the presence of other more tightly bound ligand molecules). This finding suggests that if binding of additional molecules is required for activation, it involves only a small proportion of the available receptors.

Conformational changes in the cholinergic receptor have been invoked to explain drug antagonism (the 'metaphilic effect'[28]), agonist efficacy[29,23,7], agonist-dependent mean channel conductance[30], and numerous *in vitro* observations, including those derived using fluorescence probes[1,3,6]. The advantages of using intrinsic protein fluorescence in such studies are twofold: there is a minimal perturbation of the system; and natural ligands can be used at high concentration without loss of specificity or sensitivity. The magnitudes of the effects we have observed, though small, seem to be specific. Studies now in progress are aimed at characterising the structural changes responsible for the fluorescence states of the AChR.

We thank Drs P. Adams and R. Clegg for valuable comments and Mrs A. Zechel for assistance.

R. BONNER
F. J. BARRANTES
T. M. JOVIN

Abteilung Molekulare Biologie,
Max-Planck Institut für biophysikalische Chemie,
D-34 Göttingen-Nikolausberg, FRG

1 Cohen, J. B., and Changeux, J.-P., *A. Rev. Pharmac.*, **15**, 83–103 (1975).
2 Weber, G., *et al.*, *Molec. Pharmac.*, **7**, 530–537 (1971).
3 Martinez-Carrion, M., and Raftery, M. A., *Biochem. biophys. Res. Commun.*, **55**, 1156–1164 (1973).
4 Anderson, M. J., and Cohen, M. W., *J. Physiol., Lond.*, **237**, 385–400 (1974).
5 Barrantes, F. J., Sakmann, B., Bonner, R., Eibl, H., and Jovin, T. M., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 3097–3101 (1975).
6 Grünhagen, H., and Changeux, J.-P., *C. r. hebd. Séanc. Acad. Sci., Paris*, **281**, 1047–1050 (1975).
7 Colquhoun, D., *A. Rev. Pharmac.*, **15**, 307–325 (1975).
8 Gage, P. W., *Physiol. Rev.*, **56**, 177–247 (1976).
9 Magazanik, L. G., *A. Rev. Pharmac.*, **16**, 161–175 (1976).
10 Changeux, J.-P., *Proc. sixth Int. Congr. Pharmac.*, I (edit. by Tuomisto, J., and Paasonen, M. K.), 165–177 (Forssan Kirjapaino Oy, Forssa, Helsinki, 1975).

11 Cohen, J. B., Weber, M., Huchet, M., and Changeux, J.-P., *FEBS Lett.*, **26**, 43–47 (1972).
12 Olsen, R. W., Muenier, J.-C., and Changeux, J.-P., *FEBS Lett.*, **28**, 96–100 (1972).
13 Rigler, R., Rabl, C. R., and Jovin, T. M., *Rev. Sci. Instr.*, **45**, 580–588 (1974).
14 Ellman, G. L., Courtney, K. D., Andres, V., and Featherstone, R. M., *Biochem. Pharmac.*, **7**, 88–95 (1961).
15 del Castillo, J., and Katz, B., *Proc. R. Soc.*, **B146**, 369–381 (1957).
16 Katz, B., and Miledi, R., *J. Physiol., Lond.*, **224**, 665–699 (1972).
17 Magleby, K., and Stevens, C. F., *J. Physiol., Lond.*, **223**, 173–197 (1972).
18 Bennet, M. V. L., Wurzel, M., and Grundfest, H., *J. gen. Physiol.*, **44**, 757–804 (1961).
19 Popot, J.-L., Sugiyama, H., and Changeux, J.-P., *C. r. hebd. Séanc. Acad. Sci., Paris*, D279, 1721–1724 (1974).
20 Weber, M., David-Pfeuty, T., and Changeux, J.-P., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 3443–3447 (1975).
21 Katz, B., and Thesleff, J. *Physiol., Lond.*, **138**, 63–80 (1957).
22 Rang, H. P., and Ritter, J. M., *Molec. Pharmac.*, **6**, 357–382 (1970).
23 Rang, H. P., *Br. J. Pharmac.*, **48**, 475–495 (1973).
24 Adams, P. R., *Pflügers Arch.*, **360**, 135–144 (1975).
25 Dreyer, F., and Peper, K., *Nature*, **253**, 641–643 (1975).
26 Adams, P. R., *Pflügers Arch.*, **360**, 155–164 (1975).
27 Lester, H. A., and Sheridan, R. E., *Biophys. J.*, **16**, 212a (1976).
28 Rang, H. P., and Ritter, J. M., *Molec. Pharmac.*, **5**, 394–411 (1969).
29 Barlow, R. B., Lowe, B. M., Pearson, J. D. M., Randall, H. M., and Thompson, G. M., *Molec. Pharmac.*, **7**, 357–366 (1971).
30 Colquhoun, D., Dionne, V. E., Steinbach, J. H., and Stevens, C. F., *Nature*, **253**, 204–206 (1975).
31 Jovin, T. M., in *Biochemical Fluorescence: Concepts*, 1 (edit. by Chen, R. F., and Edelhoch, H.), 305–375 (Dekker, New York, 1975).

Virtual absence of L-glutamate from the haemoplasm of arthropod blood

DURING the past decade substantial evidence has accumulated that L-glutamate is a transmitter at arthropod somatic neuromuscular junctions[1]. One of the major paradoxes in the acceptance of this evidence is the reported high concentration of L-glutamate in the haemolymph, which is sufficient on neurophysiological preparations to decrease or abolish neurally-evoked contractions[2,3]. To overcome this problem it has been suggested that either there is a barrier to L-glutamate between the neuromuscular junction and the haemolymph[3,4], or L-glutamate is bound in the haemoplasm in a pharmacologically inactive form[5] or sequestered into the haemocytes[6,7]. These suggestions would now seem to be obsolete in view of our evidence that L-glutamate is virtually absent from arthropod haemoplasm.

The neuromuscular junctions of *Lucilia sericata*, in common with other blowfly larvae, are not covered by glial cells[8,9] and so there is no cellular barrier between the synapse and the L-glutamate in the haemolymph. Injection of L-glutamate into the haemolymph of larval *L. sericata*, as found with larval *Calliphora vicina*[6], caused reversible paralysis, confirming that there is no apparent barrier to L-glutamate between the haemolymph and the neuromuscular junction in this *in vivo* preparation. Furthermore, injection of radiolabelled L-glutamate into the haemolymph reveals rapid metabolism and removal of L-glutamate to other tissues, which indicates that the L-glutamate is not bound in the haemoplasm in a pharmacologically inactive form. As fresh haemolymph has been reported to be pharmacologically inactive[5], an investigation into the electrophysiological effects of fresh haemolymph on body wall muscles 6A and 7A of *L. sericata*[10] was undertaken. These muscles are sensitive to L-glutamate which causes muscle membrane depolarisation and contraction. The effect on the membrane potential when fresh haemolymph was applied was unpredictable individually, but when repeated many times gave no significant difference between the membrane potentials before and after haemolymph application ($n = 105$). Aged haemolymph gains pharmacological activity in *Schistocerca gregaria*[6] and this is also true for aged *L. sericata* haemolymph, which evokes physiological responses similar to those elicited by L-glutamate.

Fresh haemolymph is thus pharmacologically inactive, and no barriers (physical or chemical) seem to exist between the neuromuscular junction and the haemolymph. One possible conclusion is that the haemoplasm does not contain L-glutamate and so to test this hypothesis, we analysed the haemolymph of various insects using a Locarte automatic amino acid analyser (Table 1 and Fig. 1). The haemolymph from individuals of each species was pooled, to give samples of 30 µl which were centrifuged for 4 min at 800g, the haemoplasm removed and the proteins precipitated using mainly 70% ethanol, and occasionally

Table 1 The concentration of L-glutamate and L-glutamine in the haemoplasm and haemocytes of five insects and one crustacean

	Lucilia sericata		*Calliphora erythrocephala*		*Schistocerca gregaria*		*Locusta migratoria*		*Periplaneta americana*		*Carcinus maenus*	
Place of extraction	Below spiracular sclerite				Dorsal cervical membrane				Removing hind leg		Membrane between coxa and thorax	
Animals per sample	5		3		3		3		3		1	
No. of samples	C(5)	P(6)	C(4)	P(4)	C(4)	P(8)	C(4)	P(5)	C(6)	P(7)	C(2)	P(5)
Mean L-glutamate concentration	3.3 ±1.5	— 6*t*	20.1 ±7.4	—	0.8 ±0.8*	— 2*t* 1*t*	0.4 ±0.4*	— 3*t* 1*t*	6.3 ±2.6	— 6*t*	9.6 2.6	— 3*t*
Mean L-glutamine concentration	121.3 ±50.3	1047.3 ±130.3	131.6 ±29.8	827.0 ±223.0	25.7 ±11.9	403.0 ±22.3	9.2 ±2.8	305 ±41.6	60.6 ±15.6	214.6 ±26.6	15 ±7.6	48 ±15.3

C, Haemocytes; P, haemoplasm. Units, 10^{-5} M ± s.e.
*One sample contained quantifiable L-glutamate.
t, Number of samples containing trace of L-glutamate ($< 10^{-5}$ M).

3% sulphosalicylic acid (glutamate recovery 97–99% and 95–98% respectively). The precipitate was removed by centrifugation, resuspended in distilled water, recentrifuged and the supernatants pooled. All experimental procedures were carried out at 4 °C. The results revealed that although there is L-glutamate in the haemocytes of all the species tested, it is not detectable, or only present in trace amounts ($< 10^{-5}$ M) in the haemoplasm. These results have been confirmed qualitatively by high voltage paper electrophoresis (HVPE), and a Beckman 120C amino acid analyser with an initial buffer of pH 3.25, which achieves a discrete glutamate peak away from the glutamine peak. Our data thus clearly contradict the reports of others[2-7,12-14] who found levels of L-glutamate as high as

Fig. 1 Traces taken from Locarte automatic amino acid analyser. Upper trace, Absorbance at 570 nm; lower trace, absorbance at 440 nm. *a*, Standards: glutamine, 53 nmol; glutamate, 20 nmol. *b*, Unclotted *Periplaneta americana* haemoplasm showing no detectable glutamate. *c*, Clotted *P. americana* haemoplasm showing presence of glutamate. *d* and *e*, Unclotted *P. americana* haemoplasm showing traces of glutamate ($< 10^{-5}$ M). ■, Glutamine; ▼, glutamate. All peak heights integrated by computer.

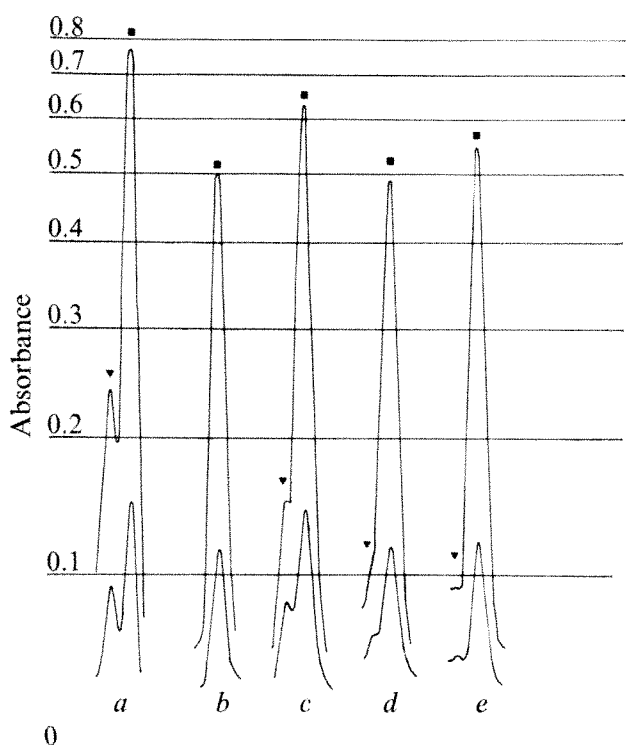5.8×10^{-4} M (ref. 5) in arthropod haemoplasm. To resolve this contradiction, we studied the methods used by these other workers, and found that these methods are capricious, for the following reasons.

(1) L-glutamine is often the most abundant amino acid in the haemoplasm (10^{-2}–10^{-3} M) and is known to be unstable[15-18] at extremes of pH. Using HVPE and radiolabelled L-glutamine, we found that the common acid protein precipitants (perchloric acid, trichloroacetic acid and sulphosalicylic acid) at the quoted concentrations are sufficient to convert some L-glutamine to L-glutamate and pyrrolidone carboxylic acid (Table 2). Furthermore, perchloric acid left at room remperature in a clear glass vessel shows a dramatic increase in hydrolysing ability (Table 2).

(2) Many published haemolymph extractions were conducted at room temperature[3,5,13]. We found it difficult to prevent the first stages of clotting at room temperature and even at 4 °C half the samples of some species had to be discarded after centrifuging because some clotting had occurred. Haemolymph analysis of these partially clotted samples revealed L-glutamate in the haemoplasm at concentrations as high as 4.5×10^{-4} M. Thus attempts to analyse partially clotted samples of haemolymph of those insects tested will reveal L-glutamate in the haemoplasm. This could explain why aged haemolymph gains pharmacological activity.

(3) Many authors did not separate haemoplasm from haemocytes[11,19,20]. Too low a centrifugation speed, however, will contaminate the haemoplasm with L-glutamate from the haemocytes. Too high a centrifugation rate on the other hand,

Table 2 Reaction products following treatment of glutamine with various protein precipitants

Treatment	Glutamate	Pyrrolidone carboxylic acid
70% Ethanol	0.3–0.8	1.2–1.7
3% Sulphosalicylic acid	2.0–4.3	2.13–5.93
15% Trichloroacetic acid	0.2–2.2	0.83–5.43
0.4 N Perchloric acid	0.7–1.0	0.83–4.62
12-h-old 0.4 N Perchloric acid	10.1–11.0	10.93–11.73
1-week-old 0.4 N Perchloric acid	14.0–18.0	13.23–17.63

Glutamate and pyrrolidone carboxylic acid are expressed as percentage produced from hydrolysis of 10^{-3} M glutamine. Incubation of glutamine with protein precipitants was for 10 min in an ice bath. The 'aged' perchloric acid had been kept at room temperature. Range given from five separate determinations. The percentages of glutamate and pyrrolidone carboxylic acid formed vary greatly and the above figures should not be taken as absolute but merely as an indication that the reaction occurs.

causes haemocyte rupture, with consequent release of L-glutamate into the haemoplasm.

(4) Other tissues have a high concentration of L-glutamate[14] which may contaminate the haemoplasm during extraction.

Thus our results show that L-glutamate is virtually absent from arthropod haemoplasm, or if detected, exists in concentrations which are far below those measured by other workers. We believe that this new finding resolves one of the major obstacles to the acceptance of L-glutamate as a neurotransmitter at arthropod somatic nerve–muscle junctions. Further details of this work will be published elsewhere.

We thank Dr J. Fox for discussion and sample analysis, Dr M. Wilkinson for use of HVPE equipment, and Dr C. Potter for helpful discussion. One of us (S.N.I.) acknowledges financial support from the SRC and Wellcome Research Laboratories.

S. N. IRVING
M. P. OSBORNE

The University of Birmingham,
Department of Zoology and Comparative Physiology,
Birmingham B25 2TT, UK

R. G. WILSON

Wellcome Research Laboratories,
Berkhamsted Hill,
Berkhamsted, Hertfordshire HP4 2QE, UK

1 Usherwood, P. N. R., and Cull-Candy, S. G., in *Insect Muscle* (edit. by Usherwood, P. N. R.), 207–280 (Academic, New York and San Francisco, 1975).
2 Murdock, L., and Chapman, G., *J. exp. Biol.*, 60, 783–794 (1974).
3 Clements, B., and May, T., *J. exp. Biol.*, 60, 673–705 (1974).
4 Faeder, I. R., and Salpeter, M. M., *J. cell. Biol.*, 46, 300–307 (1970).
5 Miller, R., Leaf, G., and Usherwood, P. N. R., *Comp. Biochem. Physiol.*, 44A, 991–996 (1973).
6 Evans, P. D., and Crossley, A. C. S., *J. exp. Biol.*, 61, 463–472 (1974).
7 Evans, P. D., *J. exp. Biol.*, 56, 501–507 (1972).
8 Hardie, J., *J. Insect Physiol.*, 22, 661–668 (1976).
9 Osborne, M. P., *J. Insect Physiol.*, 13, 827–835 (1967).
10 Crossley, A. C. S., *J. Embryol. exp. Morph.*, 14, 894–110 (1965).
11 Corrigan, J. J., and Kearns, C. W., *J. Insect Physiol.*, 9, 1–12 (1963).
12 Holden, J. S., *J. Physiol., Lond.*, 232, 61–62P (1973).
13 Stevens, T. M., *Comp. Biochem. Physiol.*, 3, 304–309 (1961).
14 Levenbook, L., *Comp. Biochem. Physiol.*, 18, 341–351 (1966).
15 Blömback, B., *Methods in Enzymology*, 11, 398–411 (1967).
16 Archibald, R. M., *Chem. Rev.*, 37, 161–208 (1945).
17 Blanshard, K. C., Bradford, H. F., Dodd, P. R., and Thomas, A. J., *Analyt. Biochem.*, 67, 233–244 (1975).
18 Perry, T. L., and Hansen, S., *Clin. chim. Acta*, 25, 53–58 (1969).
19 Benassi, C. A., Colombo, G., and Allegri, G., *Biochem. J.*, 80, 332–336 (1961).
20 Osborne, N. M., and Neuhoff, U., *Brain Res.*, 80, 251–264 (1974).

Influence of social setting on the induction of brain cyclic AMP in response to electric shock in the rat

RATS subjected to electric footshock display two disparate behaviours, Escape-attempts or aggressive attack, depending on whether the rat is shocked alone or with a conspecific[1]. Different responses in blood pressure[1], ACTH levels[2], and brainstem noradrenaline turnover[3] develop in these two patterns of behaviour. On the basis of these studies we hypothesised an additional differential response for brain cyclic AMP levels. We now report a 100% increase in whole brain cyclic AMP following shock-induced fighting but not after identical footshock given to isolated rats. This increase occurred with or without systemic pretreatment of the rats with caffeine, an inhibitor of phosphodiesterase, the enzyme which deactivates cyclic AMP.

Forty-eight male Sprague–Dawley rats (180–220 g, Simonsen, Gilroy, California) were individually housed with *ad libitum* access to food and water and then randomly assigned to one of six experimental groups. These were a control group; a saline-injected control group; a shock-alone (Escape-attempt) group; a shock-paired (Fighting) group; a caffeine-treated group, unshocked; and a caffeine-treated group which was shocked when paired (Caffeine-Fighting). After two weeks of habituation to their environment the rats were exposed to testing. Those rats which received

footshock either alone or in pairs received, within a Plexiglas enclosure, 50 2-mA footshocks of 0.4 s duration with a fixed intershock interval of 7.5 s. This procedure and apparatus have been described in detail elsewhere[4,5]. Rats did not successfully avoid the footshock and all rats, whether alone or paired, subjectively appeared to sustain comparable levels of shock as indicated by vigorous vocal and skeletal motor responses. All shocked groups were tested for 3 consecutive days in the morning.

Groups receiving drugs were injected approximately 30–60 min before behavioural testing, or at the end of the study, before killing. The saline group received 0.9% saline, intraperitoneally in a 2-ml volume, for 3 d. The caffeine groups received caffeine given as the citrate in a dose of 100 mg kg^{-1}, in a 2-ml volume, also for 3 d.

Immediately following the third day of testing and treatment, each rat was killed by microwave irradiation, in order to rapidly denature the enzymes responsible for the synthesis (adenylate cyclase) and degradation (phosphodiesterase) of cyclic AMP. The microwave apparatus used was a model 4104 Metabostat (Gerling Moore), with a power output of 5 kW at a frequency of 2,450 MHz. Animals which were shocked, either alone or paired, (with 50 footshocks during a 6.2-min period), were killed between 2–5 min after the cessation of footshock. For irradiation, rats were placed in a cylindrical animal holder. This holder allowed the rats to assume a relatively natural position and placed only the head in the microwave field. There were no difficulties with animal excitation. Irradiation for 1.5 s produced rapid and uniform heating of all brain areas. All control and caffeine-injected rats were killed in the same manner. After irradiation, brains were quickly removed from the skull, frozen on dry ice, and stored at −70 °C until analysed. Tissues were homogenised in 5% TCA, containing a known amount of tritiated cyclic AMP to monitor nucleotide recovery during purification. Cyclic AMP was assayed according to a modification of the competitive protein binding method as described by Tovey *et al.*[6] using a commercially available cyclic AMP-dependent protein kinase (Sigma).

Whole brain levels of cyclic AMP are presented in Table 1. The untreated-unshocked control group (C) had brain levels of 0.202 ± 0.011 (pmol cyclic AMP per mg of wet brain weight). The saline-injected unshocked group (S) had comparable levels of 0.206 ± 0.019. Either shock to rats alone (Escape-attempt) or the administration of caffeine to unshocked rats (CAF) elicited respective cyclic AMP levels of 0.243 ± 0.013 and 0.242 ± 0.027. Paired shocking of rats, eliciting shock-induced aggression, produced levels of cyclic AMP in the untreated (Fighting) and caffeine-treated (CAF–Fighting) groups of 0.445 ± 0.073 and 0.401 ± 0.021, respectively. A one-way analysis of variance for these data was significant at the $P < 0.005$ level ($F_{5,39} = 8.68$). Untreated shock-alone group (Escape-attempt) and the untreated shock-paired group (Fighting) significantly differed from the untreated-unshocked control group by two-tailed t test ($P < 0.05$ and $P < 0.01$ respectively). The caffeine-treated group, unshocked, did not significantly differ from the saline-injected group in this study. The shock-paired caffeine-treated group (Caffeine–Fighting) had significantly

Table 1 Levels of brain cyclic AMP (pmol per mg wet weight)

Group	N	cyclic AMP (mean ± s.e.m.)
C	7	0.202 ± 0.011
S	7	0.206 ± 0.019
CAF	7	0.242 ± 0.027
Escape-attempt	8	0.243 ± 0.013
Fighting	8	0.445 ± 0.073
CAF-Fighting	8	0.401 ± 0.021

elevated cyclic AMP when compared to the saline-injected (P <0.001) or caffeine-injected (P <0.05) unfought groups. There was no statistically significant difference between the uninjected or saline-injected control groups, or between the caffeine-treated or non-treated fighting groups. The un-injected Fighting group had markedly higher levels of cyclic AMP than the uninjected Escape-attempt group (P <0.02).

These results are in accord with other studies which have demonstrated physiological, neurohumoral, and neuro-chemical differences elicited by footshock when delivered to rats alone, inducing primarily attempted escape behaviour (a fear response) and when delivered in the presence of a conspecific, inducing fighting behaviour (an aggressive res-ponse). Williams and Eichelman demonstrated that the tail blood pressure of the rat decreased in the case of the Fighting paradigm and increased in the case of the Escape-attempt paradigm[1]. They also showed that the decrease could be selectively abolished by central destruction[7] or peripheral destruction (unpublished observations) of catecholamine neurones with 6-hydroxydopamine. Con-versely the increase of tail blood pressure in the Escape paradigm could be abolished with adrenalectomy (Williams and Eichelman, unpublished observations). Conner *et al.*[2] demonstrated that the Escape-attempt paradigm elicited greater ACTH secretion than the Fighting paradigm, con-sistent with the adrenal-blood pressure response noted by Williams and Eichelman. Stolk *et al.*[3] have reported that brainstem noradrenaline metabolism also differs in the two paradigms. Rats shocked on their own (Escape-attempt) demonstrate an increase in noradrenaline turnover in the medulla-pons during the shock period, whereas rats paired and fighting fail to demonstrate this and show an increase in noradrenaline turnover in the 1-h period after shock, which correlates with the number of attacks observed.

The enhancement of shock-induced fighting in the rat has been repeatedly associated with a facilitation of central catecholaminergic systems[8,9]. Studies in mice have shown interstrain differences in brain cyclic AMP which are genetically controlled and correlate closely with murine aggressive behaviour[10]. Additional evidence has accumulated to link cyclic AMP with central catechola-mine metabolism and behaviour[11]. The large increase in cyclic AMP with fighting is consistent with the multiple associations of aggression and cyclic AMP with central catecholamine metabolism. A potential causal and temporal relationship between the increase in cyclic AMP and in-creased noradrenaline turnover following fighting could be explored in future studies in the light of the increase in nor-adrenaline turnover following fighting observed by Stolk *et al.*[3].

The increase in cyclic AMP following footshock in the Escape paradigm should not be overlooked. It remains dwarfed, however, when compared with the magnitude of increase in the Fighting situation. Although our data in this study fail to demonstrate an increase in whole brain cyclic AMP after three daily injections of caffeine, it must be considered that an increase might have been statistically significant had a larger sample been used or a different time course for sacrifice chosen. Behavioural work[12] has shown that the greatest facilitation of shock-induced fighting following caffeine treatment occurs 4 h after injection. Caffeine given under the conditions of our study neither suppressed nor facilitated the induction of whole brain cyclic AMP; rather, the behavioural setting and the behaviour arising out of it appeared prepotent. Our findings may have further implications for understanding mam-malian coping and aggressive behaviour.

BURR EICHELMAN*
ELAINE ORENBERG
ELI SEAGRAVES
JACK BARCHAS

*Department of Psychiatry and Behavioural Sciences,
Stanford University School of Medicine,
Stanford, California 94305*

1 Williams, R. B., and Eichelman, B., *Science*, 174, 613–614 (1971).
2 Conner, R. L., Vernikos-Danellis, J., and Levine, S., *Nature*, 234, 564–566 (1971).
3 Stolk, J. M., Conner, R. L., Levine, S., and Barchas, J. D., *J. Pharmac. exp. Ther.*, 190, 193–209 (1974).
4 Eichelman, B., *J. comp. Physiol. Psychol.*, 74, 331–339 (1971).
5 Belluzi, J. D., and Grossman, S. P., *Physiol. Behav.*, 10, 133–135 (1973).
6 Tovey, K. C., Oldham, K. G., and Whelan, J. A. M., *Clin. chim. Acta*, 56, 221–234 (1974).
7 Williams, R. B., Eichelman, B., and Ng, L. K. Y., *Nature new Biol.*, 240, 276–277 (1972).
8 Eichelman, B., *Neurosci. Res.*, 5, 109–129 (1973).
9 Eichelman, B., and Thoa, N. B., *Biol. Psychiatr.*, 6, 143–164 (1973).
10 Orenberg, E. K., Renson, J., Elliott, G. R., Barchas, J. D., and Kessler, S., *Psychopharm. Commun.*, 1, 99–107 (1975).
11 Skolnick, P., and Daly, J. W., *Science*, 184, 175–177 (1974).
12 Eichelman, B., Orenberg, E., Hackley, P., and Barchas, J. D., *American College of Neuropsychopharmacology Meetings, San Juan, Puerto Rico* (1975).

Spike after-hyperpolarisation of a sympathetic neurone is calcium sensitive and is potentiated by theophylline

IN many neurones, calcium as well as sodium enters during the action potential[1]. Intracellular injection of calcium has been shown in several neurones to increase membrane con-ductance to potassium[2,3], and voltage clamp experiments have demonstrated that inward calcium currents can acti-vate outward potassium currents[4-7]. The action potential in many nerve cells is followed by a slower after-hyper-polarisation (AH), which results from increases of one or more potassium conductances[8,9]. Therefore, calcium influx during an action potential might activate part or all of the subsequent AH[2,3,10]. A calcium-dependent component of the AH has been demonstrated so far in myenteric plexus neurones[11,12] and spinal motoneurones[13]. This component of the AH is blocked by Co^{2+}, Mn^{2+} and La^{3+}, which an-tagonise calcium influx[12,13], but agents that potentiate such AHs have not been reported. We have used sucrose-gap recording[14] to investigate the sympathetic ganglion of the bullfrog, and report here that the sympathetic neurones have a calcium-sensitive component of the spike AH. Fur-thermore, this calcium-sensitive potassium conductance is potentiated by theophylline, a drug known to affect cellular calcium metabolism.

Figure 1 shows that 5 mM theophylline considerably and reversibly potentiated the action potential AH in the sym-pathetic ganglion. The AH was increased whether the action potentials were generated by orthodromic B (Fig. 1*a*) or C fibre stimulation or by antidromic stimulation of the cells (Fig. 1*b*). The predominant effect of theophylline was an enhancement of the AH duration. It increased the half-time for decay of the antidromically elicited AH up to 800%. The duration of the potentiated AH was inversely dependent on stimulus frequency, with maximal durations observed only when stimuli were administered more than 1 min apart. Theophylline enhanced the AH amplitude to a lesser extent, only increasing the peak amplitude of the antidromically elicited AH up to 150%. Additionally, theophylline often changed the AH shape from a smooth curve to one with an inflection point on the decaying slope (Fig. 2*b*) or, at times, to a biphasic W-shaped curve with early transient and late slow components. The AH reached maximum potentia-tion within 15 min of the introduction of the drug and
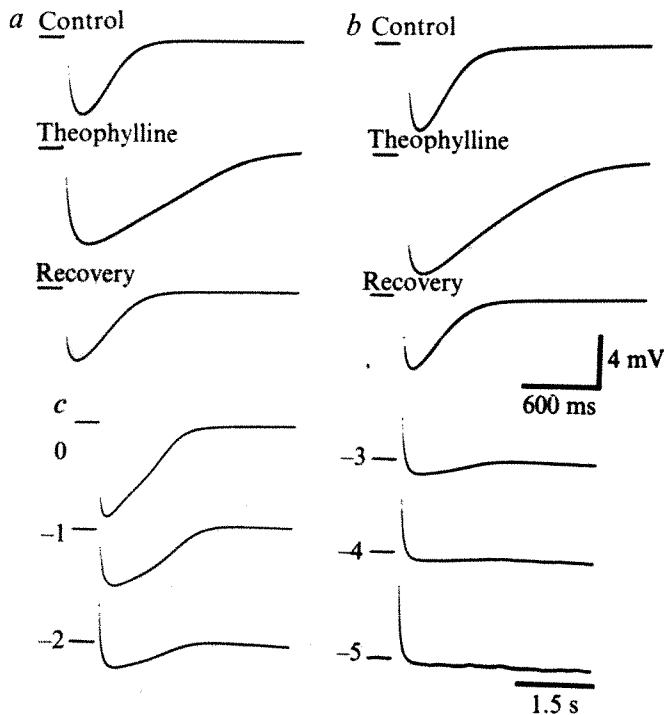
Fig. 1 Effect of theophylline on the action potential AH. *a*, Upper: AH of synaptically activated action potential. Preganglionic B fibres stimulated in the sympathetic chain rostral to the seventh ganglion with a single supramaximal stimulus. Middle: effect of theophylline on the AH 35 min after start of superfusion with Ringer solution containing 5 mM theophylline. Lower: AH 65 min after start of theophylline washout. *b*, Upper: AH of antidromically activated action potential. Postganglionic axons stimulated across the sucrose gap with a single stimulus. Site of stimulation at point where postganglionic axons exit ganglion and enter sucrose chamber. This is termed direct stimulation by Nishi and Koketsu[14]. Middle: effect of theophylline on the AH 30 min after start of superfusion with Ringer solution containing 5 mM theophylline. Lower: AH 60 min after start of theophylline washout. *c*, Electrical polarisation of theophylline-potentiated AH. Action potential elicited as in *a* and recorded 50 min after start of superfusion with Ringer solution containing 5 mM theophylline. Hyperpolarising current passed across sucrose gap[14]. Numbers to the left of each record indicate the hyperpolarising current $\times 10^{-6}$ A. Records in *a*, *b*, and *c* are from photographs of oscilloscope traces. Action potentials were lost in photographic reproduction. The experiments reported here were conducted on the tenth paravertebral ganglion of the bullfrog (*Rana catesbeiana*) by means of sucrose-gap recording[14]. The composition of the Ringer solution was (mM): NaCl, 100; KCl, 2; $CaCl_2$, 1.8; Tris-HCl, 16, pH 7.2; and glucose, 1 g l^{-1}. Calcium-free Ringer had the same composition except that $CaCl_2$ was omitted. The ganglion was continuously superfused with oxygenated Ringer solution.

returned to control after superfusion with normal Ringer for 15–30 min. Theophylline did not significantly affect the amplitude of the action potential elicited by orthodromic stimulation. Stimulus artefact obscured the antidromically elicited action potentials.

Theophylline also potentiates the slow inhibitory post-synaptic potential (slow i.p.s.p.) in both bullfrog (our work, in preparation) and rabbit[15] sympathetic ganglia. It is therefore important to compare these two hyperpolarising potentials. The slow i.p.s.p. in bullfrog ganglia is generated by a decrease in sodium conductance and thus is potentiated by moderate hyperpolarising current[16]. On the other hand, the action potential AH in frog ganglion cells has been shown to be generated by an increase in potassium conductance, and thus is decreased by hyperpolarising current[17,18]. Figure 1c shows that moderate hyperpolarisation of the resting membrane potential progressively reduced and abolished the theophylline-potentiated AH. Amplifier limitations prevented the passing of further hyperpolarising current. Furthermore, although atropine (1 μg ml^{-1}) blocks the slow

i.p.s.p.[19], it does not affect the theophylline-potentiated AH. These data distinguish the theophylline-potentiated AH from the slow IPSP and suggest that the potassium conductance during the AH is prolonged by theophylline.

The theophylline-induced potentiation of the AH was abolished by the removal of extracellular calcium. This was shown in two ways. First, as illustrated in Fig. 2b, the ganglia were superfused with a Ringer solution containing normal calcium and 5 mM theophylline until maximal AH potentiation occurred. Then they were superfused with Ringer containing theophylline but no calcium. As can be seen, the AH was markedly diminished in duration and amplitude, becoming even smaller than the pre-theophylline control. The potentiated AH returned when calcium was

Fig. 2 Effect of calcium-free Ringer solution on normal and theophylline-potentiated AH. *a*, Normal AH. Upper: AH of antidromic action potential; Middle: effect of calcium-free (0 Ca) Ringer 43 min after start of superfusion with calcium-free Ringer solution; Lower: recovery record 10 min after return to Ringer solution containing normal calcium. *b*, Effect of calcium-free Ringer on theophylline-potentiated AH. Upper: AH of antidromic spike potentiated by 5 mM theophylline; Middle: effect of calcium-free (0 Ca) Ringer containing 5 mM theophylline 33 min after start of superfusion with calcium-free Ringer solution; Lower: recovery record, 11 min after return to Ringer solution containing normal calcium and 5 mM theophylline. Records in *a* and *b* from rectilinear pen recorder (Brush Model 280). Transient depolarisation consists of stimulus artefact and action potential; it was truncated for purposes of illustration. *c*, Effect of theophylline and calcium-free Ringer on AH on expanded time scale. ————, Normal AH of antidromic spike; ...., effect of calcium-free Ringer on normal AH; — — — — —, effect of 5 mM theophylline on normal AH. Tracings of oscilloscope records. From the same experiment as *a* and *b*.
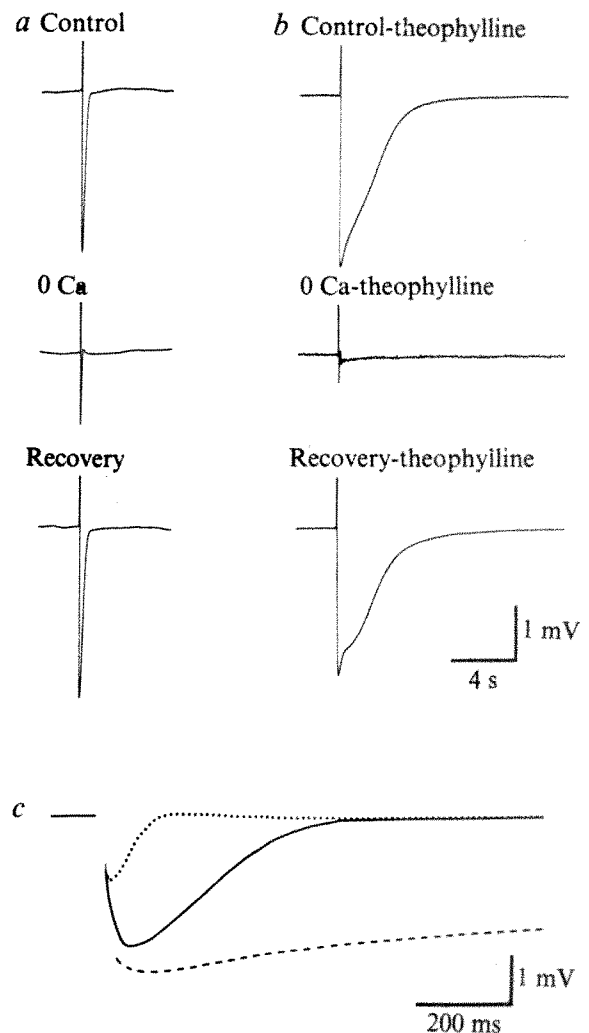
restored. Second, ganglia were superfused with calcium-free Ringer for more than 4 h. At that time the addition of 5 mM theophylline did not significantly alter the antidromically elicited AH.

Figure 2a shows that the AH elicited by antidromic stimulation in normal Ringer decreased in duration and amplitude during superfusion with calcium-free solution. This effect was reversed on readdition of calcium to the superfusate. The AH elicited in calcium-free Ringer resembled that elicited in 5 mM theophylline-calcium-free Ringer (Fig. 2b) although there was some variability in the comparative amplitudes from experiment to experiment.

Figure 2c compares the AH, on an expanded time scale, in calcium-free Ringer, in normal calcium Ringer, and in normal calcium with 5 mM theophylline, as recorded by the sucrose gap[14]. This technique records the membrane potentials of the population of ganglion cells. In experiments with intracellular recording from sympathetic ganglion cells, the removal of extracellular calcium markedly decreased the duration but not the amplitude of the AH (our unpublished observations with J. Schulman). This suggests that the decrease in AH amplitude in calcium-free Ringer recorded by the sucrose-gap technique is due to the summation of a temporally dispersed population response.

Our data show that the repolarisation phase of the AH in frog sympathetic ganglion cells depends on extracellular calcium. Koketsu and Nishi[20] demonstrated previously that calcium as well as sodium can enter these neurones during an action potential. This suggests that calcium influx during the action potential activates a potassium conductance that prolongs the duration of the AH. Alternatively, calcium influx may control the rate of inactivation of the AH potassium conductance. In the experiments reported here, we also found that theophylline greatly potentiated the calcium-sensitive potassium conductance. Theophylline may increase the stimulus-induced entry of calcium into these neurones, as it does in other tissues[21,22], thus augmenting the activation of the potassium conductance; however, we cannot exclude other possibilities, for example that theophylline enhances the calcium sensitivity of the AH potassium conductance.

Theophylline inhibits cyclic nucleotide phosphodiesterases and potentiates cyclic nucleotide-dependent processes in many tissues[23]. Cyclic nucleotides may be involved in cellular calcium metabolism[24,25] and could play a role in the augmentation of the AH by theophylline. In cardiac Purkinje fibres theophylline and cyclic AMP increase both a slow inward current carried by calcium and sodium and a slow outward current carried predominantly by potassium[22].

Calcium-sensitive potassium conductances modulate the excitability and the repetitive firing of neurones[4,10,13,26]. Theophylline may therefore prove to be a useful tool for investigating regulation of neuronal activity and responsiveness.

We thank J. A. Nathanson, J. A. Schulman and P. A. Smith for helpful discussions.

NEIL A. BUSIS
FORREST F. WEIGHT

Laboratory of Neuropharmacology,
National Institute of Mental Health,
Saint Elizabeths Hospital,
Washington, DC 20032

1 Reuter, H., *Prog. Biophys. molec. Biol.*, 26, 1–43 (1973).
2 Meech, R. W., *Comp. Biochem. Physiol.*, 42A, 493–499 (1972); *J. Physiol., Lond.*, 237, 259–277 (1974).
3 Krnjevic, K., and Liesewicz, W., *J. Physiol., Lond.*, 225, 363–390 (1972).
4 Clusin, W., Spray, D. C., and Bennett, M. V. L., *Nature*, 256, 425–427 (1975).
5 Meech, R. W., and Standen, N. B., *J. Physiol., Lond.*, 249, 211–239 (1975).
6 Mounier, Y., and Vassort, G., *J. Physiol., Lond.*, 251, 609–625 (1975).
7 Vassort, G., *J. Physiol., Lond.*, 252, 713–734 (1975).
8 Katz, B., *Nerve, Muscle, and Synapse* (McGraw-Hill, New York, 1966).
9 Barrett, J. N., and Crill, W. E., *Fedn Proc.* 31, 305 (1972).
10 Jansen, J. K. S., and Nicholls, J. G., *J. Physiol., Lond.*, 229, 635–655 (1973).
11 Nishi, S., and North, R. A., *J. Physiol., Lond.*, 231, 471–491 (1973).

12 North, R. A., *Br. J. Pharmac.*, 49, 709–711 (1973).
13 Barrett, E. F., and Barrett, J. N., *J. Physiol., Lond.*, 255, 737–774 (1976).
14 Nishi, S., and Koketsu, K., *J. Neurophysiol.*, 31, 717–728 (1968).
15 McAfee, D. A., and Greengard, P., *Science*, 178, 310–312 (1972).
16 Weight, F. F., and Padjen, A., *Brain Res.*, 55, 219–224 (1973).
17 Blackman, J. G., Ginsborg, B. L., and Ray, C., *J. Physiol., Lond.*, 167, 374–388 (1963).
18 Weight, F. F., and Votava, J., *Science*, 170, 755–758 (1970).
19 Libet, B., Chichibu, S., and Tosaka, T., *J. Neurophysiol.*, 31, 383–395 (1968).
20 Koketsu, K., and Nishi, S., *J. gen. Physiol.*, 53, 608–623 (1969).
21 Scholz, H., and de Yazikof, E., *Naunyn-Schmiedebergs Arch. Pharmak.*, 271, 374–395 (1971); Scholz, H., *ibid*, 271, 396–409 (1971); *ibid*, 271, 410–429 (1971).
22 Tsien, R. W., Giles, W., and Greengard, P., *Nature new Biol.*, 240, 181–183 (1972).
23 Robison, G. A., Butcher, R. W., and Sutherland, E. W., *Cyclic AMP* (Academic, New York, 1971).
24 Rasmussen, H., *Science*, 170, 404–412 (1970).
25 Berridge, M. J., in *Advances in Cyclic Nucleotide Research*, 6 (edit. by Greengard, P., and Robinson, G. A.), 1–98 (Raven, New York, 1975).
26 Meech, R. W., *Comp. Biochem. Physiol.*, 48A, 387–395 (1974).

Phosphorylation of myelin basic protein by vaccinia virus cores

PROTEIN kinases, catalysing the transfer of the γ phosphate from ATP to serine and threonine residues in protein substrates have been described in several enveloped viruses[1-3]. Vaccinia virus contains a protein kinase that has recently been well characterised[4]: it is located within the viral core, is not stimulated by cyclic mononucleotides and phosphorylates at least two viral acceptor proteins. The biological functions of the protein kinase and acceptor proteins are unknown, but it is possible that the viral protein kinase could also act on host proteins, thus providing a mechanism by which the virus could modify or control host processes on infection[1]. Vaccinia virus can induce a demyelinisation after experimental infection, but the mechanism for this is poorly understood[5]. It has been suggested that direct viral action on membranes may have a role in the pathogenesis of virus-induced demyelinisation[6]. We have therefore investigated the ability of vaccinia virus cores to phosphorylate human myelin membranes in an *in vitro* system and report here that they can phosphorylate purified myelin basic protein as well as basic protein in the myelin sheath.

Purified myelin basic protein was found to be an excellent substrate for the core-associated protein kinase (Figs 1 and 2a). No activator was needed, which suggests that myelin basic protein functions both as an activator and as a substrate[4]. The almost quantitative recovery of the radioactivity from the TCA experiments (Fig. 1) in the peak corresponding to the myelin basic protein on SDS gel (Fig. 2a), shows that the more than tenfold increase in ^{32}P-incorporation observed with added myelin basic protein, compared with core alone, represents phosphorylation of the myelin basic protein and not an increase in ^{32}P-incorporation in core proteins. It is possible that part of the phosphorylation of soluble myelin basic protein might actually be phosphorylation of histone contaminants. This seems rather unlikely, however, because the myelin basic protein used migrated as a single band on disc gel electrophoresis at acid *p*H and radioactivity was found only in the slices corresponding to the myelin basic protein on polyacrylamide gel electrophoresis in sodium dodecyl sulphate (SDS) (Fig. 2a). The addition of cyclic AMP or 8-bromo cyclic AMP (a cyclic AMP analogue more resistant to phosphodiesterase) did not stimulate the reaction (data not shown). Following a 1-h incubation, a total of 0.5 mol of phosphate was incorporated per mol of myelin basic protein. Determination of the isoelectric point of the myelin basic protein, by extrapolation to zero of the (*p*H-dependent) mobilities on analytical cellulose acetate sheets[11], showed a significant change in the isoelectric point from 10.8 for the native myelin basic protein, to 10.4 for the phosphorylated protein. The phosphorylated fraction was eluted from the cellulose acetate strip and contained 1 mol of phosphate per mol of myelin basic protein.

So far, the demonstration of a direct phosphorylation of host-cell components by a viral protein kinase has been hampered by the high level of endogenous protein kinase activity present in most *in vivo* or *in vitro* systems[12]. This endogenous activity

present in myelin membranes in our system can be completely inactivated by mild heating[10,13]. We have shown that such heated membranes can be used as a substrate for exogenous enzyme[10], thus allowing us to test the ability of the core-associated protein kinase to phosphorylate the myelin basic protein in an environment reflecting the native structural state. Using such heat-inactivated myelin membranes there was a significant increase in ^{32}P incorporation with added myelin membranes, compared with core alone (Fig. 1). Analysis of the phosphorylated products on SDS–polyacrylamide gel electrophoresis (Fig. 2b) shows incorporation of radioactivity over the band corresponding to the myelin basic protein. It is interesting that the viral protein kinase shows the same specificity as the endogenous protein kinase of myelin[10], because the other major myelin proteins show no incorporation of radioactivity.

We observed a much lower phosphate incorporation with myelin membranes than with isolated myelin basic protein. After incubation for 60 min, 29.7 nmol ^{32}P-phosphate per mg myelin basic protein was incorporated when isolated myelin basic protein was used as a substrate and 4.25 nmol when myelin membranes were used. Steric factors must therefore play a major part in the incorporation of ^{32}P, since our system can be compared with a solid-state reaction between the enzyme, bound to the large viral core (molecular weight $\sim 2 \times 10^9$) and the membrane-associated substrate. In the final step of our core preparation non-ionic detergents were removed, to prevent

Fig. 1 Time course of ^{32}P incorporation with core alone (▲), core and heat-inactivated myelin membranes (■) and with core and isolated myelin basic protein (●). The reaction mixture[7], with a final volume of 250 µl contained 50 mM Tris-HCl, pH 10.25, 5 mM dithiothreitol, 5 mM MgCl$_2$, 0.5 mM γ-^{32}P-ATP (30–90 c.p.m. pmol^{-1}), 30 µg of core protein and either ~100 µg of heat-inactivated myelin protein or 30 µg of purified myelin basic protein. Incubations were carried out at 34 °C. Reactions were terminated with 2.0 ml of 10% trichloroacetic acid (TCA) containing 0.06 N sulphuric acid. The material was collected on Millipore filters (0.45 µm), washed extensively, dried and counted in 10 ml of Aquasol in a liquid scintillation counter. Separate blank values, determined for different substrates, were always subtracted. Incubation of heat-inactivated myelin membranes alone for up to 1 h with the total reaction mixture gave the same value as the corresponding zero blank with added enzyme. Vaccinia virus (strain WR) and cores were prepared by slight modifications of the procedure described by Paoletti and Moss[7]. Myelin was prepared from human brain (white matter) by the method of Norton[8]. Heat inactivation was at 55 °C for 5 min and any broken membranes were removed by recentrifuging the heated membranes over a discontinuous sucrose gradient and collecting only the material banding at 0.85 M sucrose. Myelin basic protein was prepared from bovine brain[9] using Cellex-P column chromatography as a final step.
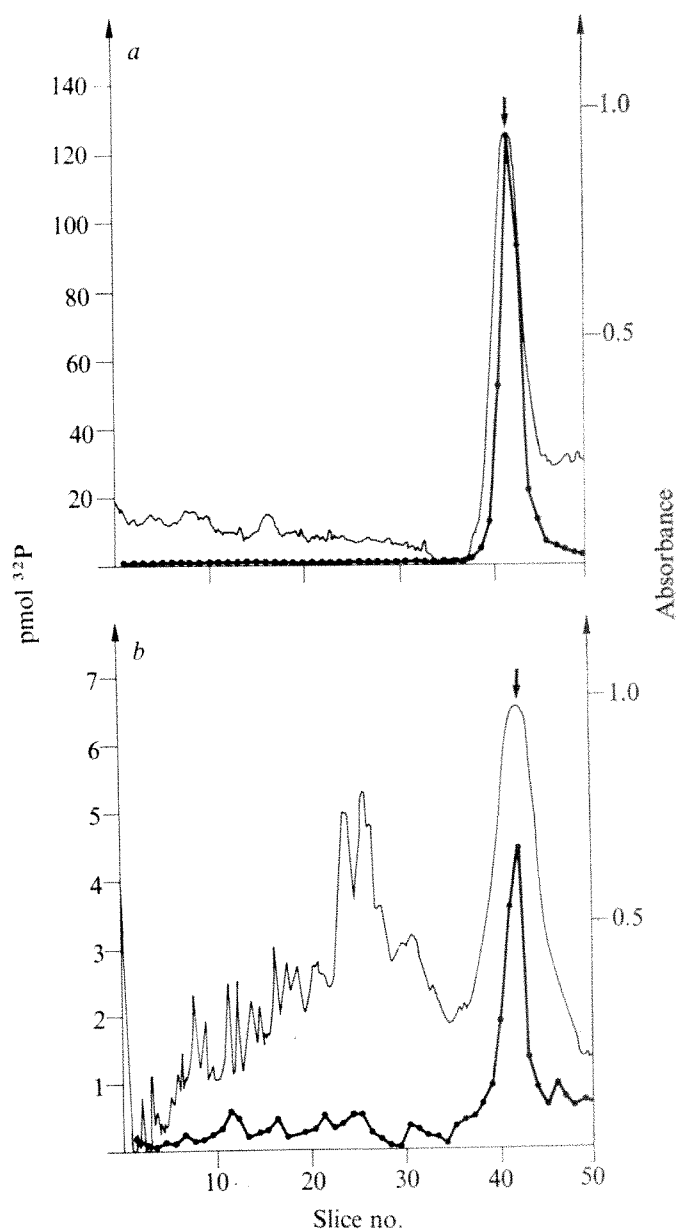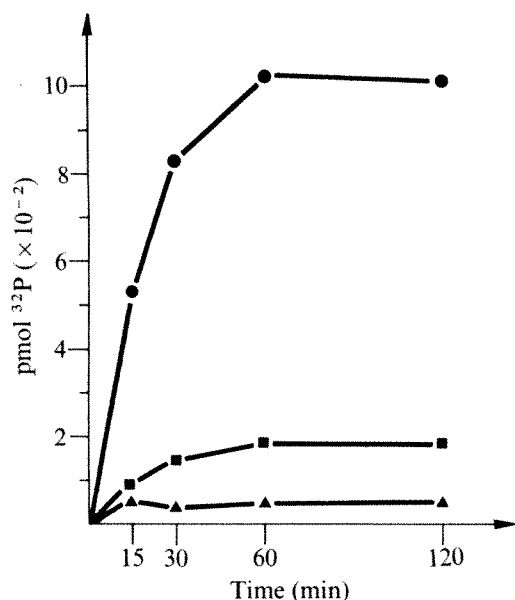

Fig. 2 Absorbance (—) of 7.5% polyacrylamide gel in 0.1% SDS and the corresponding radioactivity of ^{32}P-phosphorylated products (●). Arrow denotes myelin basic protein. The reaction mixture is as described in Fig. 1 but with a final volume of 1,000 µl; incubation was for 60 min at 34 °C. a, Incubation mixture contained 120 µg of purified myelin basic protein and 120 µg of core protein. At the end of the incubation, the mixture was cooled on ice, transferred in a centrifugation tube and the core pelleted through 0.85 M sucrose. The supernatant, containing the myelin basic protein, was collected, precipitated with TCA, washed with ethanol–water 9:1 and the resulting pellet solubilised for electrophoresis as described previously[10]. About 10 µg of protein was applied for electrophoresis. b, Incubation mixture contained 400 µg of heat-inactivated myelin membranes (prepared as described in Fig. 1) and 120 µg of core protein. After incubation the mixture was ice cooled, transferred in a centrifugation tube and the core pelleted through 0.85 M sucrose. The myelin banding at 0.85 M sucrose was collected and solubilised for electrophoresis. About 100 µg of protein was applied for electrophoresis. Preparation of gels, staining and destaining were as described previously[10]. Gels used to measure radioactivity were cut into 2-mm slices and the fractions solubilised in a toluene-based scintillation fluid containing a solubiliser (Protosol). Blank values obtained for the different substrates without enzyme were subtracted. Scanning was at 530 nm using a Zeiss spectrophotometer fitted with a linear transport accessory. The results shown are representative of three different preparations. Note the difference in the scale of the ordinate between a and b.

any solubilisation of core components. *In vivo*, however, the core dissolves during infection with the subsequent release of its components[14]. Very little is known about the fate of many of the viral enzymes, but it has been shown that virus-induced modification of host membranes is not dependent on virus multiplication, host RNA or protein synthesis[6]. It is thus tempting to speculate that phosphorylation of myelin membranes by an enzyme system that is not dependent on endogenous concentration of cyclic nucleotides may occur *in vivo*. This might have a profound effect on myelin structure and function.

As regards the biochemical mechanisms of pathogenesis, the role of the viral protein kinase may involve action on: (1) myelin basic protein, (2) nuclear histones or (3) mitochondrial cytochrome *c*, since the enzyme shows a strong proclivity for basic proteins[7] and is not specific to myelin basic protein. Whether the phosphorylation of the myelin basic protein, an encephalitogenic protein, by the vaccinia core-associated protein kinase, is relevant to the demyelinisation observed after experimental vaccinia virus infection remains to be determined.

A. J. Steck
P. Siegrist
N. Herschkowitz

Department of Pediatrics,
University of Bern,
CH-3010 Bern, Switzerland

R. Schaefer

Department of Structure Biology,
Biozentrum,
University of Basel,
CH-4056 Basel, Switzerland

1 Strand, M., and August, J. T., *Nature new Biol.*, 233, 137–140 (1971).
2 Randall, C. C., Rodgers, H. W., Downer, D. N., and Gentry, G. A., *J. Virol.*, 9, 216–222 (1972).
3 Silberstein, H., and August, J. T., *J. Virol.*, 12, 511–522 (1973).
4 Kleiman, J. H., and Moss, B., *J. biol. Chem.*, 250, 2420–2437 (1975).
5 Simon J., Hochstein-Mintzel V., Huber, H. C., Zinn, K. H., and Stickl, H., *Zbl. Bakt. Hyg.*, *I*, Abt. Orig. A232, 8–17 (1975).
6 Harter, D. H., and Choppin, P. W., in *Immunological disorders of the nervous system*, 342–355 (William and Wilkins, Baltimore, 1971).
7 Paoletti, E., and Moss, B., *J. Virol.*, 10, 417–424 (1972).
8 Norton, W. T., in *Adv. exp Med. Biol.*, 13, 327–337 (Plenum, New York, 1971).
9 Oshiro, Y., and Eylar, E. H., *Arch. Biochem. Biophys.*, 138, 392–396 (1970).
10 Steck, A. J., and Appel, S. H., *J. biol. Chem.*, 249, 5416–5420 (1974).
11 Schäfer, R., and Franklin, R. M., *FEBS Lett.*, 58, 265–268 (1975).
12 Kaerlin, M., and Horak, I., *Nature*, 259, 150–151 (1976).
13 Miyamoto, E., and Kakiuchi, S., *J. biol. Chem.*, 249, 2769–2777 (1974).
14 Fenner, F., McAuslan, B. R., Mims, C. A., Sambrook, J., and White, D. O., in *The Biology of Animal Viruses*, 207–220 (Academic, New York and London, 1974).

Ribonuclear protein formation at locus *2-48 BC* in *Drosophila hydei*

In *Drosophila hydei* salivary glands, a puff can be induced by treatment with vitamin B[6] (ref. 1) at locus *2-48 BC*. This puff is one of the series that can also be induced by temperature shock, anaerobiosis or treatments interfering with the respiratory metabolism[2]. A product of this puff seems to be a giant ribonucleoprotein (RNP) complex (diameter up to 4,000 Å) with a very specific morphology[3]. This complex consists of a core and a cortex, formed by a wreath of small particles (diameter approximately 300 Å; ref. 4). Cytochemical analysis has shown that the core of the complex is essentially free of RNA whereas the cortex does contain RNA[4]. Using phospho-tungstic acid (PTA) as a specific stain for proteins, it could also be shown that the proteins from the cortex are relatively rich in arginine and lysine residues whereas those in the core are relatively poor in such residues[5]. No similar puff product has so far been found in any other locus of *D. hydei*. In addition to the giant RNP complexes small RNP particles (diameter up to 300 Å; not to be confused with the particles from
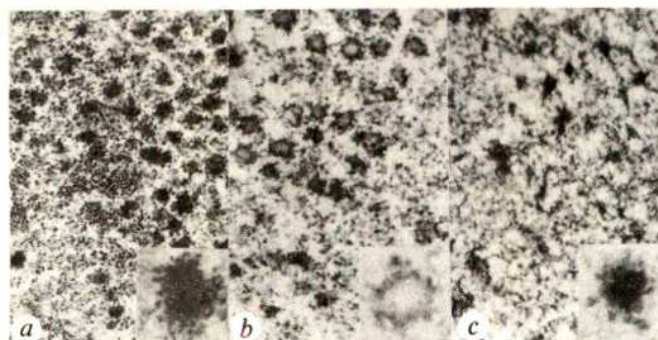

Fig. 1 Staining of RNP from locus *2-48 BC* in puff and nucleoplasm (insets). *a*, *b* and *c*, RNP particles and complexes in puff *2-48 BC*. Magnifications: *a*, *b* and *c*: ×15,360; insets: ×48,000. Glands were isolated by hand. Puffs were induced by incubation for 3 h in Poels medium containing 5×10^{-2} M vitamin B[6] (ref. 1). After fixation in 3% glutaric aldehyde in sodium cacodylate buffer (*p*H 7.2) for 30 min at 4 °C, followed by 1 h at room temperature, glands were dehydrated and embedded in Epon[8], sectioned with an LKB ultratome III and examined using a Philips EM 201. Stainings were as follows: *a*, 3% aqueous uracil acetate for 10 min and subsequently for 0.5–1 min in alkaline lead citrate[9]; *b*, after fixation and dehydration, with 3% PTA in ethanol at 0 °C overnight[10]; after washing twice with cold ethanol, glands were embedded and sectioned as described above; *c*, fixation and washing (twice) with distilled water; staining with 5% aqueous PTA (adjusted to *p*H 1 with concentrated HCl) for 3 h at room temperature (modified from Palladini *et al.*[11]), followed by three rinses with distilled water, dehydration, embedding and sectioning as described above.

the cortex) can also be seen in this puff[2,5]. A comparison of data from both ultrastructural and autoradiographical studies on RNA formation at locus *2-48 BC* has led to the conclusion that the small RNP particles present in this puff aggregate to the giant RNP complexes and thus give rise to the final puff product[6]. This conclusion is in agreement with the finding of a single RNA species in puffs obtained by microdissection of salivary glands[7]. Although the small RNP particles in the puff have not been studied extensively, it has been assumed that, based on the presence of RNA and arginine and lysine residues, the small particles form the cortex of the complex[5], leaving the origin of the cortex proteins still obscure. Here I present the results of a comparison between the staining of the small RNP particles with that of the components of the giant complexes. These studies indicate that the same components of the small RNP particle become part of the core of the complex after aggregation.

After induction of puff *2-48 BC* with 5×10^{-2} M vitamin B[6], glands were prepared for electron microscopy. A comparison was made between staining with uracil acetate–lead citrate (on grids), PTA in ethanol (on block) and aqueous PTA at *p*H 1 (on block), (Fig. 1). The results are summarised in Table 1.

It is clear that with all the three staining methods the small RNP particles stain densely. The staining with PTA was independent of the presence of RNA since neither RNase treatment (on block, followed by a mild TCA extraction), nor alkaline (0.1 M NaOH for 24 h at 25 °C) nor acid (TCA 5%, 24 h at 25 °C) hydrolysis markedly affected the staining. Thus the proteins in the small RNP particles in the puff can bind PTA in alcoholic as well as in aqueous solutions at *p*H 1. In the complex the proteins attached to the RNA in the cortex can still bind PTA in alcoholic solutions but the binding

Table 1 Staining intensity of RNP particles present in locus *2-48 BC*

Stain	Small RNP particles	Giant complexes Core	Cortex
Uracil acetate–lead citrate	+++	+++	+++
PTA, alcohol	+++	−	+++
PTA, aqueous	+++	+++	+

capacity for PTA in aqueous solutions is largely lost. The proteins in the core of the complex cannot bind PTA in alcoholic solutions, but do bind, like the small RNP particles, PTA in aqueous solution. From these data it can be inferred that, during aggregation, the small RNP particles lose an aqueous PTA-binding protein that is found later, after aggregation, in the core of the complex. The proteins in the core of the complex must therefore be assumed to derive, at least in part, from the small RNP particles. Other explanations for these observations, such as an exchange of RNP proteins with nucleoplasmic proteins, remain possible but are unnecessarily complicated. Moreover, structures resembling the core of the complex have never been observed separate from the complex. Since, in this kind of study, the available techniques do not permit quantitative analysis on the ultrastructural level, no conclusion can be drawn with respect to the amount of protein involved in the formation of the core of the complex. The presence in the complexes of nucleoplasmic proteins, besides the proteins derived from the RNP particles, cannot be excluded.

Since the giant complexes can be found in a well defined area within the puffed region[2,5], an involvement of the genome in the aggregation process could be postulated, but no evidence for such an involvement is available. So far, data concerning a possible function of the aggregation process and the proteins involved in it are also lacking. Such an analysis must await biochemical analysis of isolated complexes. It is not clear to what extent the changes in protein composition in this particular RNP product can serve as a model for RNP formation in general but the formation of this special gene product shows that RNP production can be very complicated, more so than is generally understood from the biochemical data available at this moment.

I thank Dr N. H. Lubsen for comments.

J. W. M. DERKSEN

Department of Genetics,
Faculteit der Wiskunde en Natuurwetenschappen,
Katholieke Universiteit,
Nijmegen, The Netherlands

1 Leenders, H. J., Derksen, J., Maas, P. M. J. M., and Berendes, H. D., Chromosoma, 41, 447–460 (1973).
2 Leenders, H. J., Berendes, H. D., Helmsing, P. J., Derksen, J., and Koninkx, J. F. J. G., Sub-Cell. Biochem., 3, 119–147 (1974).
3 Berendes, H. D., Symp. Soc. exp. Biol., XXV, 145–161 (Cambridge University Press, London and Cambridge, 1971).
4 Derksen, J., Willart, E., and Berendes, H. D., J. Cell Biol., 59, 661–668 (1973).
5 Derksen, J., and Willart, E., Chromosoma, 55, 57–68 (1976).
6 Berendes, H. D., in Results and Problems in Cell Differentiation, 4 (edit. by Beermann, W.), 181–207 (Springer, Berlin, 1972).
7 Bisseling, T., Berendes, H. D., and Lubsen, N. H., Cell (in the press).
8 Luft, J. H., J. Cell Biol., 25, 407–408 (1965).
9 Venable, J. H., and Coggeshall, R. J., J. Cell Biol., 25, 407–408 (1965).
10 Bloom, F. E., and Aghajanian, G. K., J. Ultrastruct. Res., 22, 361–375 (1968).
11 Palladini, G., Lauro, G., and Basile, A., Histochemie, 24, 315–321 (1970).

Long range homogeneity of physical stability in double-stranded DNA

WE present here evidence for the existence of regions with homogeneous physical stability (homostability regions) in double-stranded DNAs. Results of a study on the thermal melting of whole replicative form DNA from fd phage and of its fragments obtained by the action of restriction enzymes, strongly suggest that regional homostability seems to be an intrinsic characteristic of DNA, playing an essential part in its genetic activities.

A sequence of A–T and G–C paired bases that stabilises two nucleotide chains into one double-helical conformation is responsible for both the biological and physical characteristics of DNA; that is, the genetic information is 'written' by a variation of that sequence on the one hand, and the physical stability of the double-stranded structure is determined by the base composition on the other hand. An

indication of the second characteristic is that the melting temperature is known to increase with the G+C content. These properties of double-stranded DNA open up a possibility of probing the genetic structure by means of purely physical observations.

A series of studies on the local stability of a DNA has been carried out by a thorough measurement of hyperchromicity caused by the unfolding of the double strand with increasing temperature. Several distinctive features which have been obtained are : (1) Thermal melting is not a continuous process, but consists of a set of discrete steps (refs 1–4 and refs therein). In the case of a relatively short DNA such as λ-DNA (46,000 base pairs), it appears as several sharp peaks over a broad background in a differential display, each of which corresponds to a unit or units of melting having the same stability. (2) Spectral analysis of each peak has shown that the local G+C content in the homostability region is a major determining factor of its melting temperature (C. Akiyama, O.G., and A.W., unpublished). The Marmur–Doty relationship between the melting temperature and G+C content for whole DNA, also holds for localised melting within a DNA molecule. (3) The homostability regions also exist in long DNAs (ref.

Fig. 1 A, Melting profile of whole DNA (RF-linear) of fd phage; B, after digestion by RHinHI; that is, a mixture of HinHA+HinHB+HinHC (bold solid line); C, profiles of separated fragments, HinHA (white area), HinHB (hatched area) the and HinHC (black area), respectively. The scales of the A are same as B and C. Buffer solution: 0.1 × SSC. Doubly closed replicative form (RF) DNA of fd phage was cleaved by three different restriction enzymes (RHindII, RHinHI, RHapII), and linear RF DNA produced by RHindII, three fragments produced by RHinHI (HinHA, HinHB, HinHC), and the longest fragment produced by RHapII (HapA) were prepared as in ref. 5. Top right, the HindII cleavage site and region covered by each fragment are indicated together with an approximate location and size of fd genes[6-8]. (Roman numerals indicate the gene number.) Melting profiles were obtained with a minicomputer controlled on-lined melting plotter[3]. The displays are made in the differential form d(1−θ)/dT against T, where θ is helix content, T temperature, and (1 − θ) for the fragments is defined as (1 − θ) = (1 − θ)×(length of fragment/length of whole DNA).
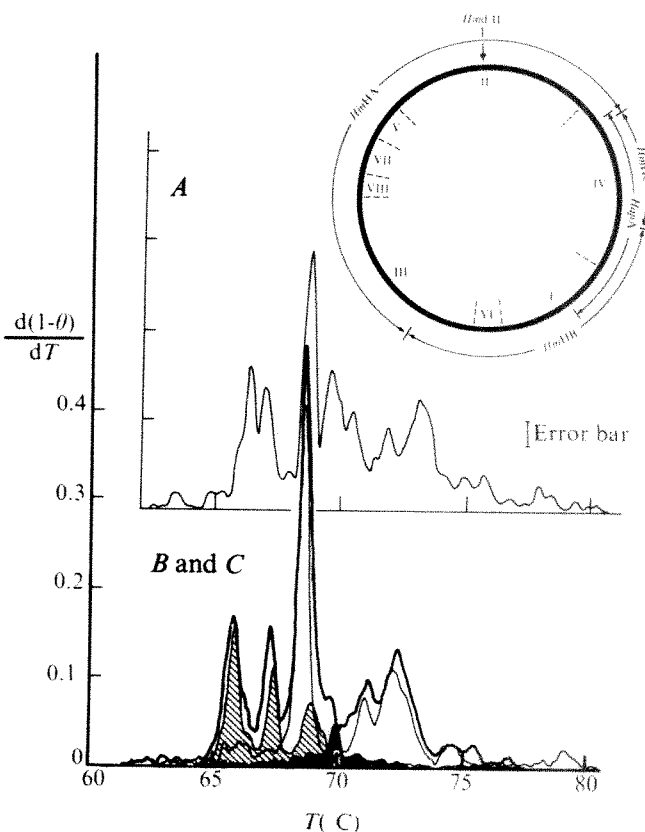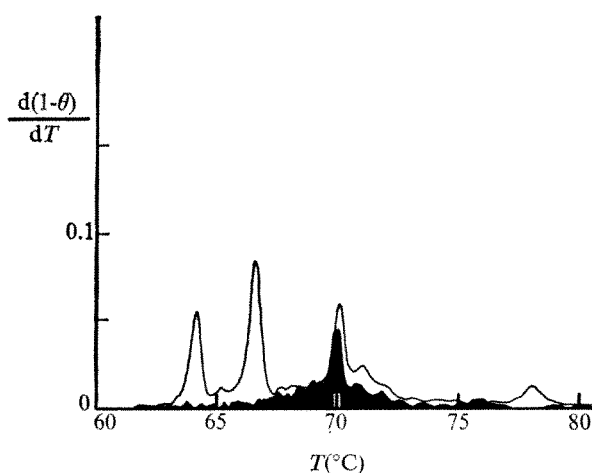
Fig. 2 Melting profiles of *HapA* (○), and of *HinHC* (●) fragment which is totally involved in *HapA*.

4 and O.G., S. Yabuki and A.W., unpublished), although the fine structures in the melting profile become obscure, probably because of random overlapping of the clustered peaks. The existence of the homostability region seems to be a general attribute of the DNA. (4) Studies on the DNA from deletion mutants have shown that a single peak in the melting profile can result from the melting of one or several homostability regions[2,3]. (5) From a computer simulation for model DNAs of given sequences[3], it is strongly suggested that a long range homogeneity in the base sequence (>500 base pairs) is necessary to produce a single sharp peak similar to those observed in our experiment. Here "homogeneity" does not mean either a regularly repeated sequence or homopolymers, but refers to a fairly long sequence consisting of over 500 base pairs in which A–T and G–C pair sequences alternate at frequent intervals made up of no more than a few base pairs[3].

In addition to the facts mentioned already, much new evidence, obtained with a much shorter DNA from fd phage and its fragments made by restriction enzymes[5], have made the characteristics of regional homostability much clearer.

The most distinctive feature of the differential melting profile of the whole fd-DNA (RF-linear; 6,000 base pairs) is that all the peaks are very sharp and well separated from one another (Fig. 1*A*). By the action of the restriction enzyme R*Hin*HI on the DNA (RF-circular), the melting profile changes to yield a set of different, but still sharp and well separated peaks (Fig. 1*B*, bold line). The individual melting profiles of each digestion product, *HinHA* (~3,500 base pairs), *HinHB* (~1,800 base pairs), and *HinHC* (~700 base pairs), were measured after separation by gel electrophoresis, and are shown together in Fig. 1*C*. The genetic map is also shown, to allow location of the fragments in it[6-8]. These three digestion product melting profiles, when summed, are found to agree with the profile of the mixed DNA (Fig. 1*B*). There are, however, not quite additive enough to reproduce that of the whole DNA (RF-linear) (Fig. 1*A*). Figure 2 shows the melting profile of *HapA*, the fragment produced by a different digestion process, compared with *HinHC*, and from the map shown in Fig. 1, *HinHC* is totally involved in *HapA*.

These experimental results from the fragments indicate that the homostability region is not uniquely defined by short range stability of the double-stranded DNA. A long range interference between the homostability regions may come through the loop entropy and electrostatic interaction of the melted region between them.

From the results, together with the genetic map, we may now conclude the following. (1) The DNA is found to consist of a number of homostability regions which come from

homogeneous base sequences consisting of 500 base pairs or more. (2) One genetic unit (the operon) may contain one or a few homostability regions. (3) Sometimes different genetic units have homostability regions with the same degree of stability. (4) A homostability region exerts the influence of the melting to other regions away from it, which sometimes produces splitting or coalescence of peaks. For instance, the highest peak of *HinHA* splits into two peaks in the profile of whole DNA, probably because of cleavage by R*Hind*II.

Location of the homostability regions on the DNA and the elucidation of the mechanism of interactions among these regions are an interesting subject for the study of the statistical mechanics of DNA chains.

Biologically, it is hard, if not impossible, to believe that such regional homostability originates in a fundamental characteristic of the genetic code itself. It is quite plausible, however, that the homostability region plays an important part somewhere in the biological process within which the DNA is closely related. If so, then the evolutionary selective force can be considered to have fixed such regions in DNA. From the size of the homostability region, recombination[9] might be one possible process which is aided by it. In any case, the wobble bases must give the necessary redundancy to make a homostability region without spoiling the biological meaning of the genetic code: the activity of proteins.

AKIYOSHI WADA
HIDEKI TACHIBANA

Department of Physics,
Faculty of Science,
The University of Tokyo,
Tokyo, Japan

OSAMU GOTOH

Saitama Cancer Center Research Institute,
Ina-machi, Saitama, Japan

MITSURU TAKANAMI

Institute for Chemical Research,
Kyoto University, Uji,
Kyoto, Japan

1 Yabuki, S., Fuke, S., and Wada, A., *J. Biochem.*, **69**, 191 (1971).
2 Yabuki, S., Gotoh, O., and Wada, A., *Biochim. biophys. Acta*, **395**, 258 (1975).
3 Gotoh, O., Husimi, Y., Yabuki, S., and Wada, A., *Biopolymers*, **15**, 655 (1976).
4 Vizard, D. L., and Ansevin, A. T., *Biochemistry*, **15**, 741 (1976).
5 Takanami, M., Okamoto, T., Sugimoto, K., and Sugisaki, H., *J. molec. Biol.*, **95**, 21–31 (1975).
6 Seeburg, P. H., and Schaller, H., *J. molec. Biol.*, **92**, 261–277 (1975).
7 Model. P., and Zinder, N. D., *J. molec. Biol.*, 231–251 (1974).
8 Horiuchi, K., Vovis, G. F., Enea, V., and Zinder, N. D., *J. molec. Biol.*, **95**, 147–165 (1975).
9 Tomizawa, J., and Anraku, N., *J. molec. Biol.*, **8**, 516 (1974); **11**, 501 (1965); **11**, 509 (1965).

Specific cleavage of chloroplast DNA from higher plants by *Eco*RI restriction nuclease

THE physicochemical properties of chloroplast (cp)–DNA in higher plants have been reported[1-4]. Cp–DNAs occur as closed circles of molecular weight 85×10^6–97×10^6. Specific identification of cp–DNAs isolated from different plants cannot, however, be carried out on the basis of buoyant density, melting points, kinetic complexity or direct length measurement by electron microscopy, since each of these techniques give values which fall within narrow limits, and only overlap with one another a little[1,2].

The analysis of restriction nuclease digests by gel electrophoresis provides a very sensitive test to compare small DNA molecule species[5]. This technique has been applied successfully to various animal mitochondrial DNAs[6] and to mt–DNA of wild and petite yeasts[7]. We show here that
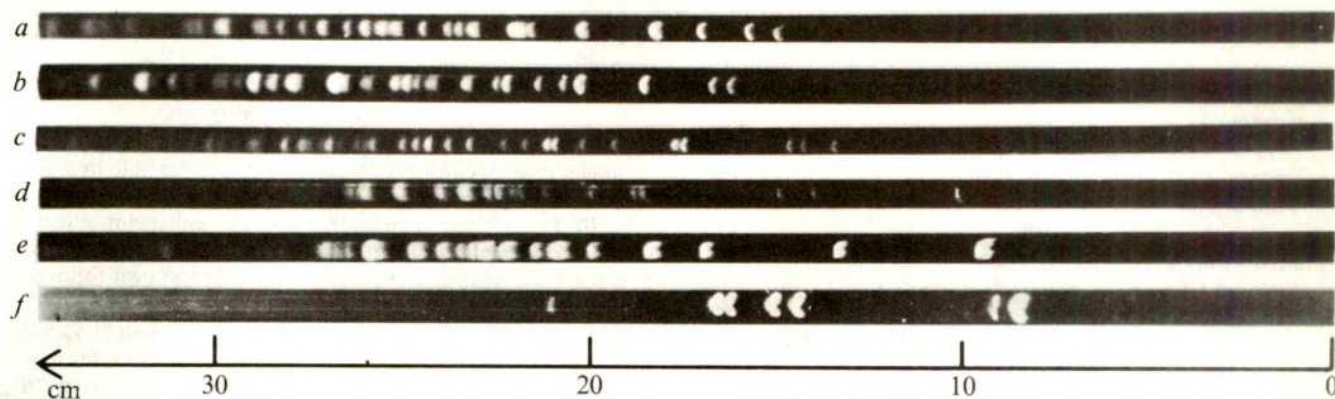
Fig. 1 Agarose gel electrophoresis of *Eco*RI digest of cp-DNAs from leaves of: *a*, pea; *b*, spinach; *c*, tobacco, *N. tabacum* var. Samsun; *d*, maize; and *e*, wheat. λ*plac*5 DNA digest (*f*) was used as a standard for molecular weight determination. The arrow indicates the direction of electrophoresis. Isolation of chloroplasts, DNA extraction from DNase-treated organelles and two-step CsCl ethidium bromide (or propidium di-iodide) gradient centrifugation of cp-DNA have been described previously[2]. The drug was removed using freshly redistilled butanol saturated with 0.01 M EDTA *p*H 8. DNA samples were then dialysed against three 500-ml changes of 0.1 M NaCl, 0.05 M Tris, 0.01 M EDTA *p*H 8.0. The cp-DNAs were pelleted at 35,000 r.p.m. for 10 h in a SW41 rotor after adjusting the NaCl concentration to 1 M. cp-DNA pellets were then dissolved in 0.1 M Tris, 0.001 M EDTA *p*H 7.5 and cleared by centrifuging for 30 s at 20,000*g*. The yields were about 30–50 μg of supercoiled cp-DNA per 100 g of leaves, depending on the plant species. DNA samples were analysed for purity and concentration by analytical CsCl gradient and were checked for circularity by electron microscopy[8]. To 20 μl of purified cp-DNA (2–3 μg DNA) were added 20 μl of 0.1 M Tris, 0.03 M MgSO$_4$ *p*H 7.5 and 5 μl of *Eco*RI endonuclease (enough enzyme to give a limit digest of 10 μg of λ DNA). Digestion was carried out at 37 °C for 3 h. The reaction was stopped by addition of EDTA to a final concentration of 15 mM and cooling to 4 °C. Each sample was treated with one volume of water-saturated phenol. Solid sucrose was added up to 20% and samples were layered on top of cylindrical 0.7% Agarose gels (type I, Sigma), 8 mm in diameter and 35 cm in length. Electrophoresis was carried out in 0.05 M Tris, 0.02 M Na-acetate, 0.002 M EDTA *p*H 8.05 (with glacial acetic acid) plus 0.18 M NaCl (ref. 5) at 100 V for 15 h, by which time, the bromophenol blue marker is ∼ 4 cm from the bottom of the gel. Gels were extruded gently from the plastic tubes and stained in darkness for 2 h in the electrophoresis buffer supplemented with 1 μg ml^{-1} ethidium bromide. They were than layered on a black background, illuminated at 254 nm with a Camag TL 900 ultraviolet lamp and photographed through an orange glass filter (Wratten 77 A) on Kodak Tri X 35-mm film. Microdol X was used for development (10 min).

cp–DNA extracted from different higher plants can be characterised and distinguished easily by gel electrophoresis of *Eco*RI digests.

Pisum sativum var. très hatif Annonay nain, *Spinacia oleracea* var. Monstrueux de Viroflay, *Nicotiana rustica* var. Pavonii, *Nicotiana tabacum* var. Samsun, *Nicotiana glauca*, *Zea mays* var. INRA 200, *Triticum aestivum* var. Capitole, were grown in a greenhouse. Chloroplasts were extracted from leaves (deribbed for spinach and tobacco) and cp–DNA isolated from DNase-treated cp-pellet according to a procedure already described[2], except that the two-step differential centrifugation used in organelle purification was repeated. The supercoiled cp–DNA was isolated as the lower ultraviolet fluorescent band of ethidium bromide (or propidium di-iodide) CsCl gradients. cp–DNAs were concentrated by pelleting after removal of drug and the purity was checked by analytical CsCl gradients (unimodal distribution). Control of circularity was achieved by Klein-schmidt spreading and positive staining with acid-alcoholic uranyl acetate[8]. The purified cp–DNAs were then digested by *Eco*RI endonuclease and the digest analysed by electrophoresis on 0.7% Agarose cylindrical gels. The *Eco*RI restriction fragments of λ*plac*5 DNA were used as mol-

ecular weight standards[5]. Gels were stained with ethidium bromide and ultraviolet fluorescence photographs were taken.

Figure 1 shows the electrophoretic pattern of the cp–DNAs isolated from pea, spinach, tobacco (*N. tabacum* var. Samsun), maize and wheat. Some 30 distinct bands were resolved in each case, about 25 moving more slowly than, or with, the bromophenol dye. The same number of bands was obtained whether electrophoresis was achieved on 35-cm or 60-cm Agarose gels. The cp–DNA of a given plant seems to be characterised by a specific digestion pattern and so may be easily distinguished from that of other plants belonging to different genera.

The characteristic band patterns of various cp–DNAs isolated in this way from several higher plants were found to be easily reproducible. For example, several independent preparations of cp–DNA have been made for each plant mentioned above. Separate digestions of these cp–DNAs were performed and the resulting band patterns found to be invariant. Identical patterns were obtained for a given plant with incubation times of cp–DNA in the presence of the *Eco*RI enzyme ranging from 1 h to 6 or 12 h.

The molecular weights of *Eco*RI fragments from

Fig. 2 Agarose gel electrophoresis of *Eco*RI digests of cp-DNAs from deribbed leaves of *a*, *Nicotiana rustica* var. Pavonii; *b*, *Nicotiana tabacum* var. Samsun; *c*, *Nicotiana glauca*; *d*, native cp-DNA of *Nicotiana tabacum* var. Samsun, a small spot (Star) of ethidium bromide was superimposed on the bromophenol blue before taking the ultraviolet photograph. Arrow indicates the direction of electrophoresis. Isolation of cp-DNAs, endonuclease digestion and electrophoresis of DNA fragments were performed as described in Fig. 1.
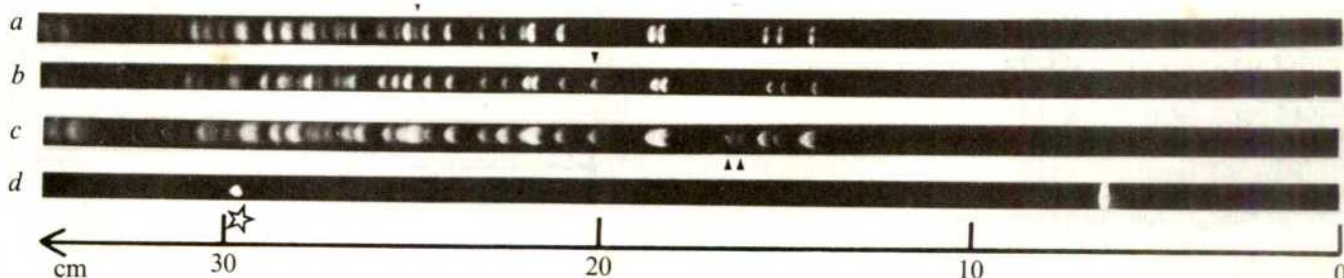
cp–DNAs shown in Fig. 1 have been determined from both their mobilities relative to λ*plac*5 *Eco*RI fragments and from their length as measured by electron microscopy after gel slicing and DNA recovery. The sum of the band sizes for each electrophoretogram of the Fig. 1 is 38×10^6–45×10^6 daltons, according to the plant, the monocots giving the lowest values. Such molecular weights represent only about half the molecular weight already estimated from the contour length of the native cp–DNA molecules and from the reassociation kinetics[1,2]; monocots displaying the lowest values here also. The fluorescence intensity of the bands does not decrease steadily with decreasing molecular weight fragments. Some bands exhibit a fluorescence intensity higher than expected and it is likely that these bands are multiple. Experiments with radioactive cp–DNA are required to state the band multiplicity precisely.

The characteristic band patterns produced after *Eco*RI digestion of cp–DNA from three plants of the genus *Nicotiana* (*N. rustica* var. Pavonii, *N. tabacum* var. Samsun, *N. glauca*) are presented on Fig. 2. It is noticeable that the three patterns are very similar, although not identical. A difference involving a few bands (indicated on Fig. 2) allows the recognition of each *Nicotiana* species. Gel electrophoresis of *Eco*RI digests provides an extremely sensitive method to evidence intergenera as well as intragenus differences at the cp–DNA level.

This method should be very useful to (1) demonstrate the function of cp–DNA by illustrating the association of genetic markers with individual groups of restriction fragments; (2) confirm the nature of parasexual hybrids; (3) analyse the phylogeny within some important genera; and (4) elucidate the origin of male sterile line.

We thank Dr H. Fukuhara for a gift of *Eco*RI endonuclease and Dr Tiollais for the λ*plac*5 DNA used in these studies.

FERNAND VEDEL
FRANCIS QUETIER
MARCEL BAYEN

Laboratoire de Biologie Moléculaire Végétale,
associé au CNRS (L.A. 40) Bât. 430,
Université Paris-Sud,
91405 ORSAY, France

1 Herrmann, R. G., Bohnert, H. J., Kowallik, K. V., and Schmitt, J. M., *Biochim. biophys. Acta*, **378**, 305–317 (1975).
2 Kolodner, R., and Tewari, K. K., *Biochim. biophys. Acta*, **402**, 372–390 (1975).
3 Kolodner, R., and Tewari, K. K., *Nature*, **256**, 708–711 (1975).
4 Kolodner, R., and Tewari, K. K., *J. biol. Chem.*, **250**, 8840–8847 (1975).
5 Helling, R. B., Goodman, H. M., and Boyer, H. W., *J. Virol.*, **14**, 1235–1244 (1974).
6 Potter, S. S., Newbold, J. E., Hutchison, C. A., III, and Edgell, M. H., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 4496–4500 (1975).
7 Morimoto, R., Lewin, A., Hsu, H. J., Rabinowitz, M., and Fukuhara, I., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 3868–3872 (1975).
8 Davis, R. W., Simon, M. N., and Davidson, N., in *Methods in Enzymology*, **21**(D) (edit. by Grossman, L., and Moldave, K.), 413–428 (Academic, New York, 1971).
9 Vedel, F., and Quétier, F., *Biochim. biophys. Acta*, **340**, 374–387 (1974).

New method of detecting singlet oxygen production

SINGLET oxygen is receiving increasing attention as a reactive species in many chemical or biological reactions (see refs 1–3 for recent reviews). We report here a new method for detecting the formation of singlet oxygen by the generation of stable nitroxide radicals from sterically hindered amines.

Excited oxygen (singlet oxygen) can be produced for example, by microwave discharge in the gaseous phase[4], by chemical reactions (the first reported one being the hypochlorite–H_2O_2

reaction[5]) or by energy transfer from triplet states of sensitisers according to the following scheme[2]

$$^3\text{Sensitiser} + {}^3O_2 \rightarrow {}^1\text{Sensitiser} + {}^1O_2$$

Singlet oxygen can be identified as an intermediate in a given reaction by (1) the quenching effect of the azide ion ($k_q = 2.2 \times 10^8$ M^{-1} s^{-1} in methanol[6]); (2) the enhancing effect of deuterated solvents[7]; (3) the identity of the oxidation products of a substrate to those obtained in a reaction known to involve singlet oxygen[3].

Quenchers other than NaN_3 have also been studied. Among these, amines constitute a separate class, probably because they are highly specific for the $^1\Delta_g$ state[3]. Typically, their quenching rate constant (k_q) ranges from 10^5–10^7 M^{-1} s^{-1} (ref. 8). The quenching mechanism has not been fully elucidated, however. The existence of a charge transfer complex between the amine and the singlet oxygen molecules has been proposed; the rationale for such an hypothesis is the good correlation found between log k_q and the ionisation energy of the amines[8]. Since such complexes are reversible, the amines are usually considered as inert quenchers[3]. A fairly new class of stable free radicals, the nitroxides, are, however, commonly synthetised by oxidation of sterically hindered amines[9]. The high specificity of amines for $^1\Delta_g$ and the possibility in some particular cases of stabilising the oxidation products, led us to try a new method in which stable nitroxide radicals would be generated by reaction with singlet oxygen and would be easily detectable by electron spin resonance (ESR) spectroscopy.

In our experiments, singlet oxygen was produced either through the classical chemical reaction between H_2O_2 and NaOCl or by energy transfer, and 2,2,6,6-tetramethylpiperidin was added to the medium. A slight ESR signal was already

Fig. 1 Nitroxide concentration against duration of light exposure ▲, Proflavin; ●, toluidine blue; ×, fluorescein; △, retinal; □, xanthotoxin; ○, without sensitiser. Sensitiser concentration, 10^{-3} M; 2,2,6,6-tetramethylpiperidin concentration, 1.15×10^{-1} M; nitroxide concentration before illumination, 0.3×10^{-4} M. Light source, Osram HBO 500; filters, Schott WG 345 and KG1; total incident intensity (300–800 nm): 14×10^{15} photons cm^{-2} s^{-1} (by chemical actinometry).
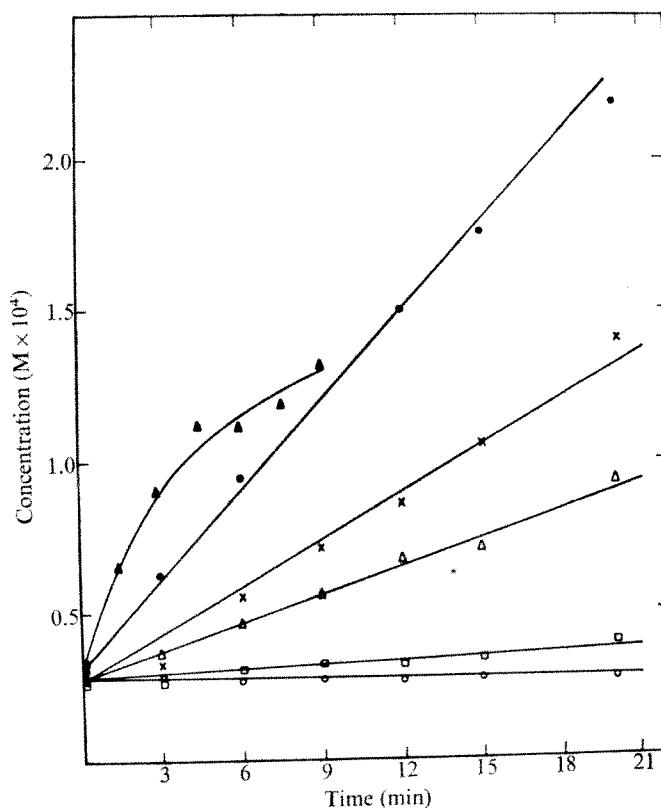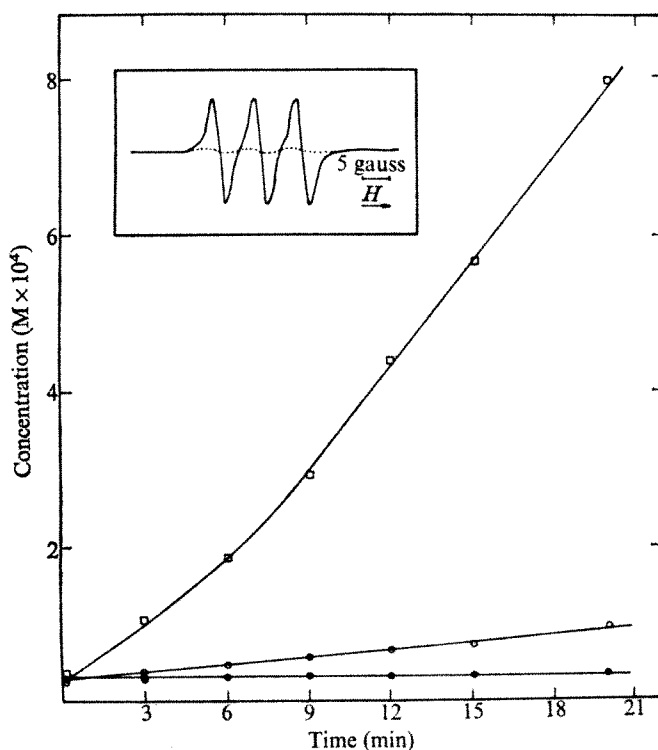
Fig. 2 Nitroxide concentration against duration of light exposure in the presence of 10^{-3} M retinal. 2,2,6,6-tetramethylpiperidin concentration 1.15×10^{-1} M; nitroxide concentration before illumination, 0.3×10^{-4} M. $\square$, CD_3CD_2OD; $\bigcirc$, CH_3CH_2OH; $\bullet$, $CH_3CH_2OH + NaN_3$ (concentration, 10^{-2} M). Illumination conditions are the same as in Fig. 1. The lag period observed in deuterated ethanol can be interpreted in terms of a buildup period for singlet oxygen steady-state concentration. Inset: ESR spectra before (. . .) and after (—) illumination. Microwave power 4 mW; modulation amplitude 1 gauss; temperature 22 °C.

observed before any reaction. This spectrum characterises nitroxide free radical[10] present as an impurity in the commercial product (Aldrich); its concentration does not exceed 0.3‰.

When the chemical reaction is performed in the presence of the amine, free radical production occurs. The corresponding ESR spectrum is identical to the one observed before reaction but its intensity is greatly increased. This suggests that it is the presence of singlet oxygen which leads to nitroxide formation.

For testing the photochemical process, various sensitisers were studied. Toluidine blue and fluorescein are known to generate $^1O_2$ (refs 11 and 12). It has only been suggested that retinal and proflavin are able to do so[13,14]. On the other hand, although there are some indications[15] that furocoumarins do not react with oxygen in biological sensitisations Poppe and Grossweiner recently[16] attributed lysozyme photodynamic inactivation to singlet oxygen generated by energy transfer from xanthotoxin triplet state; this latter compound was therefore studied.

Photochemical experiments were performed mainly in ethanol. One millilitre of sensitiser solution (1 mM) was introduced into rectangular quartz cuvettes (1×1 cm) and 20 µl of amine were added. Aliquots were taken in calibrated capillaries (Drummond microcaps 50 µl) before exposure to light and after different exposure times. Samples were immediately measured with a Varian E3 ESR spectrometer.

Illumination of the different chromophores also led to nitroxide production. Their rate of formation is shown in Fig. 1. In our experimental conditions, proflavin gave rise to the steepest initial slope but underwent rapid auto-oxidation. Toluidine blue, on the other hand, gave a fairly linear yield up to a 20-min exposure. Therefore it seems to be a very good sensitiser for nitroxide photoproduction, comparable to

retinal and fluorescein, although less effective. With respect to xanthotoxin, the small but measurable yield observed seems to confirm the results reported by Poppe and Grossweiner[16].

To prove definitively that singlet oxygen is implicated, the effects of sodium azide and deuterated ethanol were studied and the results are presented in Fig. 2. In deuterated ethanol, the free radical rate formation increases more than tenfold whereas in the presence of azide, free radical production is abolished.

When experiments were carried out in nitrogen-bubbled solutions, no nitroxide formation occurred. The illumination was usually carried out in an open cell in equilibrium with the atmosphere. In those conditions, the ESR linewidths are fairly large (2.25 gauss) due to the high oxygen concentration in ethanol (6.4 mM). When experiments were performed in a sealed cell, an important linewidth reduction was observed during the photochemical reaction. This phenomenon reflects oxygen consumption within the medium.

In conclusion, a new method for detecting singlet oxygen production can be derived from the experiments reported here. The results shown in Fig. 2 indicate that the reaction is highly specific for singlet oxygen. The technique allows a specific study of free radical formation without interference from other possible side products. Moreover the method, which is convenient and easily set up, can detect free radicals at concentrations as low as 100 nM. Preliminary results performed in aqueous solutions gave similar results to those obtained in ethanol. Possible applications of the method to biological reactions have still to be explored.

Y. LION
M. DELMELLE
A. VAN DE VORST

Département de Physique Atomique et Moléculaire,
Université de Liège, 4000 Sart-Tilman, Belgium

1 Foote, C. S., Science, 162, 963–970 (1968).
2 Foote, C. S., Acc. Chem. Res., 1, 104–110 (1968).
3 Kearns, D. R., Chem. Rev., 71, 395–427 (1971).
4 Noxon, J. F., Can. J. Phys., 39, 1110–1119 (1961).
5 Khan, A. V., and Kasha, M., Nature, 204, 241–243 (1964).
6 Hasty, N., Merkel, P. B., Radlick, P., and Kearns, D. R., Tetrahedron Lett., 1, 49–52 (1972).
7 Merkel, P. B., and Kearns, D. R., J. Am. chem. Soc., 94, 7244–7253 (1972).
8 Ogryzlo, E. A., and Tang, C. W., J. Am. chem. Soc., 92, 5034–5036 (1970).
9 Rozantsev, E. G., Free Nitroxyl Radicals (edit. by Ulrich, H.) (Plenum, London and New York, 1970).
10 McConnel, H. M., and McFarland, B. G., Q. Rev. Biophys., 3, 91–136 (1970).
11 Pottier, R., Bonneau, R., and Joussot-Dubien, J., Photochem. Photobiol., 22, 59–61 (1975).
12 Livingston, R., in Autoxidation and Antioxidants (edit. by Lundberg, W. O.), 1, 249 (Wiley Interscience, New York, 1961).
13 Politzer, I. R., Griffin, G. W., and Laseter, J. L., Chem.-Biol. Interact., 3, 73–93 (1971).
14 Saito, I., Inove, K., and Matsuura, T., Photochem. Photobiol., 21, 27–30 (1975).
15 Musajo, L., and Rodighiero, G., in Photophysiology, 7 (edit. by Giese, A. C.), 115–147 (Academic, New York, 1972).
16 Poppe, W., and Grossweiner, L. I., Photochem. Photobiol., 22, 217–219 (1975).

Triplet state involvement in primary photochemistry of photosynthetic photosystem II

THE primary photochemical reaction of photosynthesis is widely accepted to be a one-electron transfer from a chlorophyll (or bacteriochlorophyll) species to an acceptor entity[1]. Until recently it had been assumed that the excited state precursor to the electron-transfer reaction was a singlet state in each of the photosynthetic photosystems. This view is based on the fact that estimated fluorescence lifetimes[2] of $\sim$ 7 ps at the reaction centre are much too short compared with intersystem crossing times of the order of 5 ns for chlorophyll in vitro[3]. Robinson[3] pointed out, however, that the increase in fluorescence yield observed when the photosystem II (PS II) trap is closed is far less than that expected if photochemistry, such as electron transfer, were the sole quenching process. He and others[4,5] have proposed the involvement of a triplet state in the mechanism of the photochemical electron transfer, but this

suggestion has received little attention. Blankenship *et al.*[6] reported an electron spin resonance (ESR) signal arising from photosystem I (PS I) which initially exhibited microwave emission due to chemically induced dynamic electron polarisation (CIDEP). They maintained that this observation indicated that the excited state precursor to the formation of the radical species must have been in a triplet state to account for the CIDEP phenomenon. We report here the first observation of ESR signals exhibiting the CIDEP phenomenon from free radicals associated with PS II of chloroplasts and whole algae. As discussed below, this provides additional evidence that triplet states may be involved as intermediates in the mechanism of photochemical electron transfer in photosynthesis.

In our experiments we used a Varian E12 ESR spectrometer with the addition of a flash-photolysis excitation and detection system[7], a Photochemical Research Associates Model 610 flash system with a flash half life of ~ 25 μs and a Biomation Model 610B Transient Recorder buffered into a Fabritek Model 1072 Instrument Computer. The flashes were filtered by a Corning CS2–64 red filter. We modified the response time of the ESR spectrometer so that the rise time was 50 μs using 100-kHz field modulation and have observed kinetic responses at temperatures from 10 K to 300 K in the whole algal systems.

Figure 1*a* shows a typical kinetic response for *Scenedesmus obliquus* algae grown on a 99% deuterated medium[8]. We have also obtained responses from spinach chloroplasts, but signals from the whole algae *Anacystis nidulans* and *Scenedesmus obliquus* were more intense and more reversible than signals obtained from the chloroplasts. A negative excursion in Fig. 1*a* corresponds to the microwave emission mode and a positive excursion to the microwave absorption mode of ESR. If the photochemistry of PS II is blocked, by means of several red flashes before freezing the algae in the presence of 0.1 mM hydroxylamine and 20 μM DCMU (3-(3,4-dichlorophenyl)-1,1-dimethylurea)[9], the kinetic response is greatly diminished (Fig. 1*b*). This, and further evidence given below, leads us to believe that the detected ESR signals arise from free radicals associated with PS II. The signal shown in Fig. 1*a* was reproducible from the same sample for many thousands of light flashes.

ESR spectra at a fixed time after the flash were obtained from kinetic traces taken at many magnetic field positions close to $g = 2$. It was apparent in spectra from normal algae grown on protonated media that two separate fast transients gave rise to superimposed spectra in the free electron region. Furthermore,

Fig. 1 A typical flash-induced ESR kinetic profile for fully deuterated *Scenedesmus obliquus* frozen as a thick suspension in 20 mM phosphate buffer (pH = 7.5) with 10 mM NaCl and 50% glycerol at 10 K with 0.1 mW microwave power. Collection of 128 flashes in the presence of 20 μM DCMU and 100 μM hydroxylamine: *a*, after freezing in the dark; *b*, after freezing following several red flashes at 300 K.
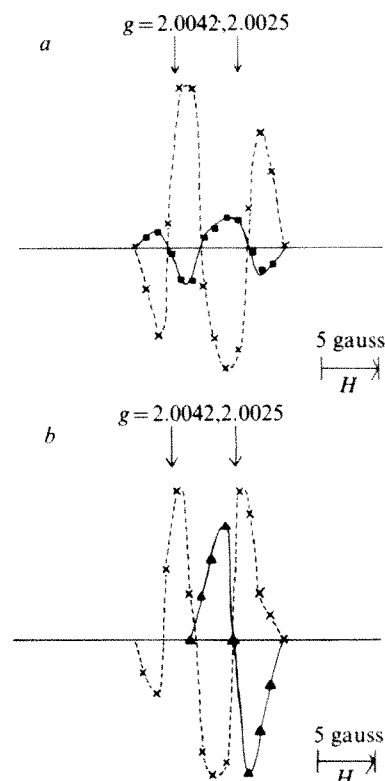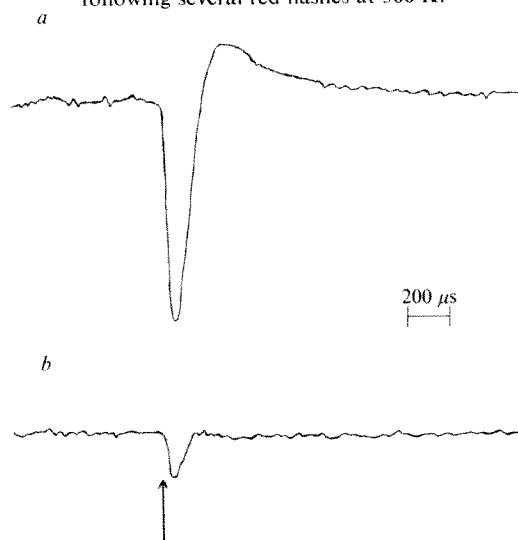

Fig. 2 *a*, A point-by-point magnetic field profile of the rapid transient amplitudes (×) and of the slow transient amplitudes (■) induced by flashing 256 times fully deuterated *Scenedesmus obliquus* suspended in 20 mM phosphate buffer (pH = 7.5) with 10 mM NaCl and 50% glycerol at 100 K with 1 mW microwave power and 2.5 gauss modulation amplitude at 100 kHz. The *g* factors are accurate to a precision of ±0.0003. *b*, A point-by-point field profile of the rapid transient amplitudes (×) and of transient amplitudes of Signal 1 (▲) induced by flashing 1,024 times fully deuterated *Scenedesmus obliquus* suspended in the same medium at 200 K with the same instrumental conditions and precision as in *a*.

the observed spectral line shapes for the fast transients were mixtures of absorption and dispersion signals arising, we believe, from rapid passage effects[10]. The spectra of rapid and slow transients (Fig. 2*a*) for deuterated algae, however, were found to be simpler because of less overlapping of the spectra of the separate transients as a result of the collapse of some unresolved proton hyperfine structure. Rapid passage effects are also responsible for some distortion of the line shapes in Fig. 2*a*.

Figure 2*b* shows the field profile of the fast transients for deuterated algae relative to the reversible part of P700 (signal 1) at a temperature of 200 K. It is evident that the fast transient signal is in opposite phase to that of signal 1, but nonetheless, it is readily detectable at this temperature. It is significant that the new transient is detectable at temperatures down to 10 K, where signal 1 is not significantly reversible for detection on a millisecond time scale. This behaviour implies that the reversible transient signal is independent of the primary photochemistry of PS I.

We interpret the kinetic response in Fig. 1*a* to fast and slow decay processes in opposite phase. The fast decay corresponds to electron spin lattice relaxation causing the spin populations (initially inverted) to relax to a Boltzmann distribution. The slow decay corresponds to a chemical disappearance of the radicals. On the basis of the spectra observed for the rapid transients from both protonated and deuterated algae, we can assign the two observed radical species, one with $g = 2.0042$ and another with $g = 2.0025$. The latter corresponds very well to the ESR spectrum of P680[+], the chlorophyll *a* species acting as the primary donor of PS II and the former could arise from a

plastosemiquinone anion radical. There is now good evidence for P680 being the primary electron donor[11-24] in PS II, and there is good optical evidence that a plastoquinone molecule acts as the primary acceptor[21-26], although no one to date has detected an ESR signal from the primary acceptor. The slow phase of decay of Figs 1*a* and 2*a* could correspond to an electron-transfer back reaction between the primary donor and acceptor. A decay half time of ~ 1 ms for both the donor and the acceptor is comparable with the 3-ms decay measured optically at low temperatures in chloroplasts[21-24].

Given that our interpretation of the kinetic response and ESR spectra is correct, we should explain how a population inversion of spin states might be established when the free radicals are created by the photochemical electron transfer from P680 to the primary acceptor. The phenomenon of CIDEP was first observed[27] by Fessenden and Schuler and has since been observed for many radicals formed in solution[28]. As yet, the work of Blankenship *et al.*[6] and our present work are the only reports of CIDEP from a biological solid-state system.

Two mechanisms have been proposed[28] to explain the observation of CIDEP. The first is the radical pair mechanism, but it can be excluded in the case of a solid-state system, because it requires free diffusion of radicals in a dynamic fluid state. The photochemical triplet mechanism involves an intersystem crossing into a triplet excited state with unequal rates of populating the three triplet levels. If electron transfer occurs before spin-lattice relaxation in the triplet state can restore a Boltzmann equilibrium, then both the resulting radicals should be observed with initially inverted spin populations. Clearly, the triplet mechanism can explain our results since we have a solid-state system.

If a triplet state of the P680 chlorophyll species is the precursor to photochemical electron transfer in PS II, then intersystem crossing must occur very quickly. Fluorescence decay measurements[29] indicate that fluorescence $t_\frac{1}{2}$ are no more than ~ 100 ps in PS II, which suggests that there must be some mechanism for enhancement of the intersystem crossing rate to allow for such rapid quenching of fluorescence. One possibility is that the manganese present in PS II is in close proximity to P680: transition metal ions are known to enhance intersystem crossing rates[30]. One other possibility is that P680 is a dimer of chlorophyll *a* which has a structure analogous to that of P700 in PS I. The quenching of the excited singlet state would then correspond to a charge separation within the dimer pair (P^+P^-). In such a charge-transfer state, the separation between singlet and triplet states would be small (because of the greater distance between the two unpaired electrons), and so an enhanced intersystem crossing rate would be expected.

Our results do not prove that a triplet state must be an intermediate in the photochemistry of PS II, as there may be another mechanism which could produce the CIDEP phenomenon from a single-state precursor, but they should provide impetus to a search for evidence of short lived triplet states in photosynthetic systems, using picosecond optical spectrometers.

ALAN R. MCINTOSH
JAMES R. BOLTON

Photochemistry Unit,
University of Western Ontario,
London, Ontario, N6A 5B7, Canada

1 Sauer, K., in *Bioenergetics of Photosynthesis* (edit. by Govindjee), 116–181 (Academic, New York, 1975).
2 Zankel, K. L., Reed, D. W., and Clayton, R. K., *Proc. natn. Acad. Sci. U.S.A.*, 61, 1243–1249 (1968).
3 Robinson, G. W., *Brookhaven Symp. Biol.*, 19, 16–48 (1967).
4 Robinson, G. W., *Proc. natn. Acad. Sci. U.S.A.*, 49, 521–529 (1963).
5 Franck, J., and Rosenberg, J. L., *J. theor. Biol.*, 7, 276–301 (1964).
6 Blankenship, R., McGuire, A., and Sauer, K., *Proc. natn. Acad. Sci. U.S.A.*, 72, 4943–4947 (1975).
7 Warden, J. T., and Bolton, J. R., in *Creation and Detection of the Excited State*, 2 (edit. by Ware, W. R.), 63–97 (Marcel Dekker Inc., New York, 1974).
8 Daboll, H. F., Crespi, H. L., and Katz, J. J., *Biotech. Bioengng*, 4, 281–297 (1962).
9 Vermeglio, A., and Mathis, P., *Biochim. biophys. Acta*, 368, 9–17 (1974).

10 Mailer, C., and Taylor, C. P. S., *Biochim. biophys. Acta*, 322, 195–203 (1973).
11 Malkin, R., and Bearden, A. J., *Proc. natn. Acad. Sci. U.S.A.*, 70, 294–297 (1973).
12 Malkin, R., and Bearden, A. J., *Biochim. biophys. Acta*, 325, 266–274 (1973).
13 Butler, W. L., *Accts Chem. Res.*, 6, 177–184 (1973).
14 Kitajima, M., and Butler, W. L., *Biochim. biophys. Acta*, 325, 558–564 (1973).
15 Visser, J. W. M., and Rijgersberg, C. P., *Proc. Third Int. Cong. Photosyn. thesis* (edit. by Avron, M.), 399–408 (Elsevier, Amsterdam, 1974).
16 Van Gorkom, H. J., Tamminga, J. J., and Haveman, J., *Biochim. biophys. Acta*, 347, 417–438 (1974).
17 Knaff, D. B., and Malkin, R., *Biochim. biophys. Acta*, 347, 395–403 (1974).
18 Ke, B., Sahu, S., Shaw, W., and Beinert, H., *Biochim. biophys. Acta*, 347, 36–48 (1974).
19 Malkin, R., and Bearden, A. J., *Biochim. biophys. Acta*, 396, 250–259 (1975).
20 Amesz, J., and de Grooth, B. G., *Biochim. biophys. Acta*, 376, 298–307 (1975).
21 Mathis, P., and Vermeglio, A., *Biochim. biophys. Acta*, 368, 130–134 (1974).
22 Mathis, P., and Vermeglio, A., *Biochem. biophys. Res. Commun.*, 56, 682–688 (1974).
23 Mathis, P., and Vermeglio, A., *Biochim. biophys. Acta*, 396, 371–381 (1975).
24 Haveman, J., Mathis, P., and Vermeglio, A., *FEBS Lett.*, 58, 259–261 (1975).
25 Van Gorkom, H. J., *Biochim. biophys. Acta*, 347, 439–442 (1974).
26 Pulles, P. M. J., Kerkhof, P. L. M., and Amesz, J., *FEBS Lett.*, 47, 143–145 (1974).
27 Fessenden, R. W., and Schuler, R. H., *J. chem. Phys.*, 39, 2147–2195 (1963).
28 Wan, J. K. S., Wong, S. K., and Hutchinson, D. A., *Accts Chem. Res.*, 7, 58–64 (1974).
29 Beddard, G. S., Porter, G., Tredwell, C. J., and Barber, J., *Nature*, 258, 166–168 (1975).
30 Calvert, J. G., and Pitts, J. N., *Photochemistry*, 317–321 (Wiley, New York, 1966).

Fluorescence lifetimes of haem proteins excited into the tryptophan absorption band with synchrotron radiation

FLUORESCENT spectra of tryptophan residues in proteins are a powerful tool in the study of the intraprotein medium, primarily as an indicator of local conformational changes of the residue microenvironment. Fluorescence quantum yields of tryptophan in the case of haem proteins, however, are very small $(Q \leq 0.002)^1$, and a dramatic increase $(Q' \sim 0.2)$ is observed after haem removal. The photodissociation of carbon monoxymyoglobin, when irradiated using ultraviolet light, has indicated the transfer of excitation from the aromatic amino acid to the haem group[2,3] following a mechanism of the Förster type[4]; this transfer has a probability[1] at least 100 times greater than the fluorescence, and therefore some 20 times greater than the other competing radiation-free processes. Consequently lifetimes have been considered to be extremely small and no attempts have therefore been made to measure them. This type of information is, however, vital for our understanding of the quenching mechanism by haem. We have therefore reinvestigated the problem with respect to decay times for ox, horse and adult human haemoglobins and their α and β subunits when excited in the tryptophan absorption band. Our data are not refined—lifetime values reported here have a rather poor precision (~ ± 0.25 ns)—but those obtained using ox, horse and human samples demonstrate similar behaviour. In some selected cases data from haemoglobin and apohaemoglobin have been compared.

Membrane-free haemoglobin samples were prepared by 22,000 r.p.m. centrifugation[5]. The haemoglobin solutions were (poly)crystallised according to Drabkin's method[6], and fresh samples were used. Haem-free haemoglobin samples were obtained from human adult haemoglobin by precipitation in acid acetone medium and then renatured as described previously[7].

Haemoglobin splitting in corresponding subunits was carried out according to Bucci's procedure[8]; thiol group regeneration, following Geraci[9]; and free SH titration in accordance with Boyer[10]. The purity and the homogeneity of α_{SH} and β_{SH} chains was controlled by electrophoretic analysis over starch gel, and by optical absorption spectroscopy. Their non-denatured character was verified by determining their capacity of association for haemoglobin reconstitution.

Ferrihaemoglobin was prepared by oxidation of deoxyhaemoglobin *in vacuo* with potassium ferricyanide in moderate (5–10%) excess (pH ~ 7), the ferricyanide surplus being eliminated by gel filtration through a G25 Sephadex column in phosphate medium.

For all samples a common solvent was used: 10^{-2} M KH_2PO_4–Na_2HPO_4 (pH = 7) buffer in sterile and clear commercial Biosedra water, presenting a weak fluorescence. The pre-existing organic phosphates (DPG) in the haemolysate have not been eliminated, their concentration in the stock solution being estimated[11] at between 0.7 and 1 equivalent per tetramer of

protein. Care was taken to avoid bacterial contamination (filtration through adequately treated 0.2-μm Millipore).

Native protein solutions (about 2.10^{-6} M, of haem or of polypeptide chain) from three different donors of each mammalian species (four in the human case) have been studied at 25 ± 0.5 °C and 3 ± 0.5 °C, always with fresh samples not older than 15 h, and stored at 4 °C in the above buffer.

Fluorescence decay times were obtained using the synchrotron radiation from the Orsay Storage Ring (ACO; for a technical description of the ACO machine, see Fourth, Fifth and Sixth int. Conf. High Energy Accel., Dubna (1963), Frascati (1965) and Cambridge, Massachusetts (1967) and ref. 12.) as a pulsed light source coupled with the Single Photoelectron Timing Method[13,14]. The ACO machine was working in the e^+, e^- collision ring mode for a high energy physics experiment with 510-MeV electron energy and 25–10-mA electron circulating current; in these 'parasitic' conditions the available light intensity is smaller by a factor of 5–10, compared with the photon flux obtained using ACO working in the Storage Ring mode, fully geared to synchrotron radiation.

Samples were excited at three different wavelengths (270, 280 and 290 nm), selected using a high flux vertical 'Lavollée-type' monochromator[15], the excitation spectral bandwidth being about 32 Å. All the emitted fluorescence with wavelength higher than 340 nm, was detected at right angles through a set of three Schott-type filters (WG 320, 330 and 345) using a RTC 56 DUVP/03 photomultiplier. Decay curves were recorded by counting signals, in the positive mode for a given time, from the excited haemoprotein solution; and then, in the negative mode, signals for an equivalent active time from the buffer solvent at the same excitation wavelength. With this latter operating mode (which is justified by the fact that synchrotron radiation is perfectly constant in shape and amplitude), and because of the small optical density of the haemoprotein solutions (see below), the stray light level from the excitation is minimised. Therefore, when verified using a MgO diffusing screen, it turns out to be less than 1%, in any case.

We will comment here only on some unexpected results (unpublished). We have roughly verified the fluorescence efficiency depletion when going from apohaemoglobins to haemoglobins, by comparing photon-counting rates in the same experimental conditions (globin counting rate $\sim 4 \times 10^4$ Hz; ferrihaemoglobin, $\sim 5 \times 10^2$ Hz, for example); lifetimes are, however, very little changed (the main component in the ferrihaemoglobin decay is only $\sim 20\%$ smaller: τ(globin) ~ 4.3 ns, τ(ferrihaemoglobin) ~ 3.5 ns, both at 3 °C. (Apohaemoprotein and haemoprotein fluorescence decay are, in most cases, non-exponential, with the main component usually the shortest one.) Because our data are not refined, we cannot find significant changes in fluorescence decays using the three excitation wavelengths mentioned above.

To eliminate the possibility that the observed emission was not due to haem losses in the haemoprotein solution[21], some samples were controlled by comparing the optical absorption in the 280-nm and 540-nm regions of their ferricyanomet derivatives. These controls have proved that the total amount of haem-free globin, if any, could not exceed 0.5%; a haem loss artefact can therefore be excluded.

An important decrease in the quantum yield, when fluorescence lifetimes are only slightly changed would indicate some sort of quenching with a strong 'static' component[16,17], the de-excitation being related to some non-fluorescent complex[18] between the fluorophore (tryptophan) and the inhibitor (haem). It may be, however, that this non-fluorescent complex is not permanent, the static quenching indicating statistical fluctuations; the oscillations of tryptophan residues (it is known that the tryptophan emission in the globin is strongly depolarised, (unpublished results)) could correlate periodically, however, with the collective movements of the macromolecules in a particular orientation, and thus escape from the influence of haem. We could then imagine that, statistically, a small fraction ($\sim 1\%$) of the total excited tryptophan population are emitting their characteristic fluorescence. The small, but significant decrease in the ferrihaemoglobin (tetramer)

decay time, compared with that of the globin (dimer), could be explained by significant changes in the tryptophan microenvironment, reflecting global conformational changes induced in the macromolecule by the presence of haem.

On the $\alpha_{SH}O_2$ chain, besides the static quenching by the haem groups, there must also be a 'dynamic' one, because of the marked decrease both in the counting rate and in the lifetime: τ(α globin) ~ 4.4 ns, τ($\alpha_{SH}O_2$) ~ 1.0 ns; both at 3 °C (ref. 16). The $\beta_{SH}O_2$ chain shows a clear non-exponential decay, the main component having a lifetime of the order of 1.4 ns at 25 °C; the other one, much weaker, is much shorter (probably of the same order of magnitude as that of the α_{SH} chain, see below). The same behaviour is found in the case of ferrohaemoglobin with an O_2 ligand; here the main component of the decay is about the same (~ 1.5 ns; always at 25 °C) but the contribution of the short component is much weaker than in the β chain. Note, however, that in the isolated α chain (monomer) there is a single tryptophan (tryp 14) residue; whereas in each subunit of the β protein (tetramer), there are two tryptophan residues (tryp 15 and 37). In addition, very little is known about the conformation of isolated α and β chains. Nevertheless there is a fundamental difference between the three haem proteins—$\alpha_{SH}O_2$, $\beta_{SH}O_2$ and HbO_2—and the ferrihaemoglobin, that is, the absence of O_2 ligand in the latter.

We have also observed that lifetimes decrease substantially when increasing the temperature from 3 to 25 °C: in the ferrihaemoglobin τ changes from ~ 3.5 ns to ~ 2.4 ns; from ~ 1.0 ns to ~ 0.5 ns, for the $\alpha_{SH}O_2$ chain. This time, in contrast with the unexpected results mentioned above, counting rates change in the same direction and in the same proportion, as would be expected. This behaviour could indicate substantial changes in the macromolecular conformation (it is known that the protein structure is very sensitive to temperature).

In addition, we must also consider a possible effect from oxygen; indeed, one could be tempted to attribute, at least partially, the dynamic quenching to oxygen diffusing into the protein[18,19]. Although the macroscopic concentration of dissolved oxygen in the aqueous samples does not change very much in our temperature range ($\sim 3.9 \times 10^{-4}$ M at 3 °C; $\sim 2.5 \times 10^{-4}$ M, at 25 °C), it is possible to imagine the presence of important local fluctuations: the intraproteic medium, covered with hydrophobic residues[20], could, in some conditions, act as an 'oxygen trap', the Henry laws being no longer valid. Another alternative hypothesis would be to consider the oxygen diffusion as being influenced by temperature-induced changes in the protein structure.

Further work will be necessary to answer some of these many questions—precise fluorescence quantum yield measurements, emission frequency distribution, time-resolved frequency distribution, and fluorescence time-dependent depolarisation.

We thank the DMA (the High Energy Physics Experiment at ACO) engineers and physicists for support, and Drs A. Tramer and L. Lindqvist for discussions.

B. ALPERT

UER Biomédicate des Saints-Pères ELM,
Université René Descartes, 45 rue des Saints Pères,
75270 Cedex, Paris 06, France

R. LOPEZ-DELGADO

Laboratoire de Photophysique Moléculaire,
Université de Paris Sud, Bâtiment 213,
91405 Orsay, France

1. Weber, G., and Teale, F. J. W., *Discuss. Farad. Soc.*, **27**, 134 (1959).
2. Bücher, T., and Kaspers, J., *Naturwissenschaften*, **33**, 93 (1946).
3. Bücher, T., *Adv. Enzym.*, **14**, 1 (1953).
4. Förster, T., *Ann. Physik*, **2**, 55 (1948).
5. Perutz, M. F., *J. Cryst. Growth*, **2**, 54 (1968).
6. Drabkin, D., *J. biol. Chem.*, **164**, 703 (1946).
7. Rossi-Fanelli, A., Antonini, E., and Caputo, R., *Biochim. biophys. Acta*, **30**, 608 (1958).
8. Bucci, E., and Fronticelli, C., *J. biol. Chem.*, **240**, 551 (1965).
9. Geraci, G., Parkhurst, L. J., and Gibson, Q. H., *J. biol. Chem.*, **244**, 4664 (1969).
10. Boyer, P. D., *J. Am. chem. Soc.*, **76**, 4331 (1954).
11. Ames, B. N., *Meth. Enzym.*, **8**, 115 (1968).
12. Dagneau, *et al.* (collectif of LURE), *Ann. de Phys.*, **9**, 9 (1975).
13. Lopez-Delgado, R., Tramer, A., and Munro, I. H., *Chem. Phys.*, **5**, 72 (1974).
14. Lopez-Delgado, R., *Proc. second SSRP Users Group Meeting*, Stanford (1975).
15. Lavollée, M., Lopez-Delgado, R., and Robin, S., *Rev. Sci. Instr.* (in the presss).
16. Vaughan, W. M., and Weber, G., *Biochemistry*, **9**, 464 (1970).
17. Perrin, F., *Ann. Phys.*, **12**, 169 (1929).
18. Lakowicz, J. R., and Weber, G., *Biochemistry*, **12**, 4161 (1973).
19. Alpert, B., and Lindqvist, L., *Science*, **187**, 836 (1974).
20. Perutz, M. F., Muirhead, H., Cox, J. M., and Goaman, L. C. G., *Nature*, **219**, 131 (1968).
21. Rossi-Fanelli, A., and Antonini, E., *J. biol. Chem.*, **235**, no. 2, PC5 (1960).

reviews

Images of terrestial processes

Peter J. Smith

The Bowels of the Earth. By John Elder. Pp. vii+222. (Oxford University: London, 1976.) £6.00.

ACADEMIC writers have usually sought to develop as impersonal a style as possible, presumably in the belief that to do otherwise would be to impair their reputation for impartiality and objectivity or detract from the science itself. This tradition rightly persists, although at the same time there is a minor (but growing) and not entirely unwelcome movement towards the injection of 'personality' into the business. As far as the Earth sciences are concerned, the latter trend seems to have originated with Peter Wyllie (*The Dynamic Earth*, Wiley, 1971) who began in a small way by placing more overt emphasis on the scientists involved and later extended the technique with great effect in *The Way the Earth Works* (Wiley, 1976).

There is no doubt that in moderation, and performed with Wyllie's flair, an attempt to present a more sparkling view of science can be entertaining and give a more realistic picture of how science actually develops. But as Elder demonstrates here, it can also degenerate into irrelevance, absurdity and even crudity. The fact that the title of his book has achieved recognition in the new (1976) edition of *The Concise Oxford Dictionary* hardly disguises the vulgarity and gimmickry of its use in a serious context. It is not even accurate. How a word which refers metaphorically to 'innermost parts' can be taken to apply to the whole Earth from core to crust (not to mention the Earth as a body in space) is beyond imagining.

None of this would matter perhaps if it were not indicative of similar errors of judgement throughout the book, chiefly in the addiction to inane metaphor and analogy. The Earth, for example, is variously a jam-pot, a stone, a jumping bean, an egg and much else besides, presumably in the mistaken impression that these homely analogies will appeal to what in these days of sexual equality Elder chooses to call 'the man in the street'. All of which is a pity because it obscures what is at heart a fascinating conception, namely, that the Earth may be viewed as a series of dynamic processes

occurring on a wide range of scales. To adapt one of his more sensible metaphors, Elder has given us an interesting sketchbook of the Earth—a collection of interacting models and images of terrestrial processes which vary in size, content, validity and credibility in reflection of what we know and can deduce from a body which is in the last resort inaccessible.

It is an unusual, if at times idiosyncratic, gallery which could be viewed with benefit by all Earth scientists, not least the 'experts' whom Elder explicitly excludes. But whether, as he fondly imagines, it is equally accessible to the

man in the street, I take leave to doubt; such an unfortunate creature would find himself bombarded by 'rather a lot of mathematical formulae' and is unlikely to find complete consolation in the index-*cum*-glossary. Some background in science would seem to me to be prerequisite. But with that reservation, and not forgetting the irritations and imperfections, I would recommend the book as an important way of looking at the terrestrial machine. □

Peter J. Smith is a Reader in Earth Sciences at the Open University, Milton Keynes, UK.

Molecular orbital theory

The Determination and Interpretation of Molecular Wave Functions. (Cambridge Monographs in Physical Chemistry, Vol. 3.) By Erich Steiner. Pp. viii+205. (Cambridge University: Cambridge and London, July 1976.) £11.50.

THERE are so many books on molecular orbital theory that it is hard to find a gap in the literature. One unfilled need is for a book at an elementary level which will bridge the separation between theoretical and experimental chemists. In particular such a book should enable the experimentalist to use the wave function computer programs which are so freely available thanks to the Quantum Chemistry Program Exchange. This book only fulfills this need in part.

The level is elementary and reasonably up to date, but the book contains no more than is to be found in a couple of chapters of standard texts supplemented by one or two recent review articles. Only molecular orbital theory is included. Indeed the preface states that "since 1960 there has been a declining importance of the valence bond approach as a practical tool of the computational quantum chemist"; a view which would not be supported by Goddard or the Italian or Bristol schools.

The early chapters summarise standard theory in a clear and concise manner, and introduce the required notation. Further chapters are devoted to the orbital approximation and be-

yond; to the representation of the orbitals and charge densities with a final look at the chemical bond. It is hard to quarrel with the content although the form in which it is presented reflects the author's view of the subject. Others might feel that natural orbitals, for example, are of interest less in a chapter on electron distribution than in the section on correlation energy.

Within the constraints of the book one definite omission is a detailed look at open-shell methods. So much published work is restricted to closed-shell ground states of molecules, possibly because a convenient account of the problems and techniques involved when there are unpaired electrons is not to be found in most simple texts. There is no coverage of molecular properties other than charge distribution.

For experimentalists an additional chapter on how to produce data for a standard computer program would have been helpful, together with some examples of how to enter these data and interpret the output. Without this, the book should be valuable as a reference work for the conveted who run calculations but probably not sufficient to enable the practical chemist to run his own. On the other hand, armed with this book he should appreciate what the theoretician has done.

Graham Richards

Graham Richards is a lecturer in the Department of Physical Chemistry and a Fellow of Brasenose College, Oxford, UK

Access to soil ecology

The Distribution and Diversity of Soil Fauna. By John A. Wallwork. Pp. xii + 355. (Academic: London and New York, April 1976.) £11; $17.25.

IN 1970 Wallwork published his *Ecology of Soil Animals* which has since been accepted as an invaluable reference and teaching aid. With commendable frankness he introduces the new book as "something of an Adam's rib, a companion volume . . .". To anyone who has shared the development of his subject over the past thirty years, reading the new work is a most enjoyable experience for many original insights, a clear and often amusing style, and some superb illustrations. The atmosphere conveyed is almost that of a set of personal discussions, perhaps following a series of seminars, among people already familiar with the jargon and many of the ideas.

The chapters are mainly devoted to major ecosystems in turn, grasslands, moorlands, forest soils and so on. Here Wallwork encounters the organisational problem which faces all writers of ecology books: the material to be covered is related in many logical 'dimensions' but must be discussed in linear order. Many topics, common to different ecosystems must either therefore be repeated or introduced rather arbitrarily in a particular chapter. The second practice is followed—for example, discussions of pedology and feeding biology in the grassland chapter—and the problem of re-locating a particular discussion is not fully met by means of cross references. One feels the need for a diagrammatic master plan in the introduction. The excellent index, however, goes a long way to overcome this difficulty.

The chapter on "agricultural practice and the soil fauna" is a short but valuable introduction to the literature of a fast-expanding field; and the final chapter faces the perennial problem of how so diverse a fauna persists in an apparently rather uniform and narrow spatial zone. Throughout, the discussion is immensely strengthened by Wallwork's experience of soils from all parts of the world.

This is hardly, as claimed on the book jacket, "a basic introduction to the subject for students unfamiliar with the soil fauna". Many terms and ideas are either used without definition or are only defined after first introduction (the vocabulary of saprophage, fungivore, and so on, and the terms 'stentotopic' as against 'eurytopic'). These are perhaps further indications of the true readership of the book: it is probably best thought of as a synthesis, a companion volume and extension of *The Ecology of Soil Animals* and a valuable means of access to the literature of soil ecology.

Amyan Macfadyen

Amyan Macfadyen is Professor in the School of Biological and Environmental Sciences at the New University of Ulster, Coleraine, Northern Ireland.

Somatic chromosomes

Chromosomes in Mitosis and Interphase. (Handbuch der Mikroskopischen Anatomie des Menschen.) By Hans George Schwarzacher. Pp. viii + 182. (Springer: Berlin and New York, 1976.) DM 136; $55.80.

THIS monograph is essentially concerned with the structure and organisation of the somatic chromosomes of man and I must say at once that I enjoyed reading it. Schwarzacher's excuse for writing such a book is the veritable flood of new information on human chromosomes that has been unleashed following the development and application of a wide variety of new techniques over the past few years.

The information and discussion is parcelled into ten chapters beginning with simple accounts of chromosome morphology, nomenclature and the human karyotype. There is a good review of what little we know of chromosome structure in the interphase nucleus as revealed by histochemical and electron microscope studies, supplemented by the observations on prematurely condensed chromosomes obtained by the method of Rao and Johnston using the cell fusion technique. Heterochromatin and chromosome fine structure, as studied by electron microscopy, are considered in detail, as are the main techniques utilised to produce the various banding patterns in human chromosomes. The two chapters on chromosome banding and fine structure are particularly

⬤ *Climate and the Environment* by J. F. Griffiths (UK edition by Paul Elek reviewed in *Nature*, **262**, August 5, 523, 1976) is also published in the USA by Westview Press, Boulder, Colorado 80301. Price : $12.75.

good, and present a clear account of the facts and the various interpretations and speculations on mechanisms and structure. In such a rapidly developing field it is inevitable that some parts of the text are a little out of date—for example, some of the more recent work on human satellite DNAs is not discussed, but this is not a major criticism.

Geneticists who have not been following closely the human cytogenetic scene should be impressed by a wealth of chromosome polymorphisms that have been unearthed in the chromosome complements of various human populations using modern cytological techniques, and these are admirably summarised by the author. Man's chromosomes are varied indeed and, for example, our own studies in Edinburgh reveal that some 3% of the local population have inversions involving the whole or part of the C band on chromosome 1; that all of us are heteromorphic for one or more C- or Q-band regions; and that the chromosome phenotype of an individual may be almost as distinctive as his fingerprints! Schwarzacher, in fact, points out that an analysis of chromosome polymorphisms would give successful paternity exclusion tests in about 75% or more of cases.

The text of the book is very clearly written; it is well illustrated and easy to read with only rare glimpses of minor problems of translation from the German, as in sentences such as "With this term the morphologic separation of complete chromosome sets of genomes is meant."

The general style tends to be rather more descriptive than discursive and I personally would have preferred more discussion of some of the interesting points brought out by the author. For example, a photograph and two lines of text impart the information—substantiated by all workers interested in chromosome replication—that daughter strands are always placed to the outside of parental templates. This implies a certain pattern of DNA segregation which must of itself reflect on organisation of chromosome structure—something that is not really considered in the book.

This little book is timely, and shows that we have come quite a long way in the past five years or so towards revealing the enormous diversity and variation in structure of man's chromosomes, but we have a long way yet to travel before we can really understand what these variations mean both in terms of chromosome organisation and in possible consequences to the individual. Swarzacher's summary sets a very useful background and is to be highly recommended to those who are already journeying in this field or who are about to start on their travels.

H. John Evans

Professor Evans is Director of the Medical Research Council Clinical and Population Cytogenetics Unit, Western General Hospital, Edinburgh, UK.

obituary

Dr A. P. Rowe, the wartime Head of the Telecommunications Research Establishment—the home of radar—died at Malvern in May 1976 at the age of 78.

Albert Percival Rowe trained as a meteorologist and worked in the Air Ministry's Laboratory in Imperial College after the 1914–18 War. From there he became Personal Assistant to H. E. Wimperis, the first Director of Scientific Research at the Air Ministry; and in June 1934—alarmed by the inadequacy of air defence in Britain, as he related in his book *One Story of Radar* (1948)—he warned the Air Ministry that "Unless science evolved some new method of aiding air defence, we were likely to lose the next war if it started within ten years". In the following November, Wimperis proposed the formation of a Committee for the Scientific Survey of Air Defence, which quickly took shape as the famous Tizard Committee.

The story has been told many times of how the principle of radar was outlined by Watson-Watt to the Committee at its first meeting, as a result of a query about death rays put to him by Wimperis, and how the independent approach of F. A. Lindemann and Churchill led to intense friction. Rowe's part in these events was to be the first Secretary of the Tizard Committee, and to decide with Wimperis to buy Bawdsey Manor, just north of Felixstowe, to house the research and development work that would be needed before radar could be made operational. In the spring of 1938 Rowe succeeded Watson-Watt as Superintendent of what was then known as the Air Ministry Research Establishment, Bawdsey, and he remained Superintendent, later Chief Superintendent, of this Establishment and its successors until the end of the war, by which time it was the great Telecommunications Research Establishment at Malvern with a staff of 3,000.

The basic notion of radar came from a Government establishment, the Radio Research Station of the National Physical Laboratory, staffed by men who—with Rowe—had gone into Government service between the wars rather than into universities. Great though the contributions of the latter were subsequently to be, they would mainly have come too late for the Battle of Britain, and it was our national good fortune that there was enough talent in Government service to

enable us to prepare for that critical phase. Recruiting from the universities had started in 1935, with men such as E. G. Bowen, R. Hanbury Brown, and A. G. Touch; and it was these recruits who, together with Watson-Watt's original nucleus of L. H. Bainbridge Bell and A. F. Wilkins, did the essential early work. Academic reluctance to work on military projects was dispelled by the Anschluss of 1938, and by the outbreak of war there was a flood of willing — and distinguished — scientific recruits, most of whom went into radar.

Rowe therefore found himself at the head of a great body of talent, from which ideas and inventions flowed throughout the war. Somehow he had to guide his establishment of scientists and engineers, few of them accustomed to Service or Civil Service procedures, into effective relationships both with the Royal Air Force and with industry. Throughout the war he held vigorously the belief that the work his Establishment was doing was vital; and at the end of it he wrote "We believe that without radar Fighter Command and the 'famous few' would have lost the Battle of Britain. We believed that without radar the night attacks by enemy bombers would have devastated our industries. We believed that without radar the work and gallantry of Bomber Command would have been largely wasted; and we believed that without radar the sea-war would have been lost and that there would have been no invasion of Europe". Although by 1948 he had begun to wonder whether this 'overwhelming belief' had been entirely justified, it had been his guiding star throughout the war and he would declare it to his staff and outsiders alike.

He was completely sincere, and it was his sincerity that enabled him to keep control of what could easily have been a very unruly Establishment. Sometimes this same sincerity led him into slightly bizarre situations : on one occasion during the war he sent a circular letter to the wives of his staff telling them that their husbands were doing work of the utmost national importance but that, unlike the serving officers with whom they were working, they were entitled only to ordinary civilian rations. Since there were various deficiencies in wartime food, it was essential that each wife should do her best to supplement her husband's diet, notably in respect of vitamins and

vitamin B in particular. This vitamin was to be had in abundance in 'Bemax' and so would she please feed this to her husband every morning. In a similar vein he had issued a station instruction at Bawdsey pointing out that the dispersal of the several research teams in huts hundreds of yards apart inhibited communication and the cross-fertilisation of ideas: he therefore urged all members of the staff to buy second-hand bicycles, and added the address of a cycle dealer in Felixstowe whom he had ascertained to have an adequate supply. But despite the comic effect produced by such instructions, he was obviously so genuine that he kept the respect, which with the years developed into affection, of his staff.

Perhaps his greatest individual contribution was the 'Sunday Soviet' held weekly in his room which, after the move of TRE to Malvern College, was of course the Headmaster's Study. Sundays were convenient for such men as Lord Cherwell, Sir Robert Renwick, and senior serving officers to visit Malvern, and Rowe would gather members of his staff for a free-for-all discussion which would last from morning well into the afternoon. Bernard Lovell, then one of Rowe's men, has recorded that "A Commander-in-Chief once complained in Rowe's office that the most lowly assistant in the Establishment would not turn a screw unless he could be acquainted with the latest strategic situation in the Middle East". The interplay between scientists and serving officers had not been started by Rowe; from their experiences in the 1914–18 War, Lindemann and Tizard, for all their differences, were united in the importance of this interplay, but it reached one of its highest expressions in Rowe's Soviets. These meetings were one of the main factors in the ultimate superiority of British radar over its counterpart in Germany, where scientists and military men were not nearly in such effective contact.

The successes of TRE were many. As regards actual inventiveness, the ideas of course originated with individuals, to whom full credit should be given; and the characteristic problems associated with getting individuals to work in a corporate body are probably less in war than in peace, where the importance and urgency of a common aim are less clear. But, even so, with all the brilliance of individuals and the corporate spirit engendered by national emergency, TRE would have been far

less effective if it had not been so well guided, and this was Rowe's contribution. He would certainly have been blamed had TRE failed, and in its success he deserved a greater honour than the CBE that was awarded to him in 1942. In fact, the whole of the radar effort was miserably recognised during the war: the Americans, who were glad to have a contingent at TRE, gave Rowe a more fitting recognition—the Medal for Merit, their highest civilian decoration, awarded directly by the President.

Writing to me in 1974 of such matters he said "I have never written of those things to anyone till now. I feel strongly about two lines of a prayer, 'To toil and not to seek for rest, To labour and not to seek for any reward'. Certainly I never thought of it during the war, when vast numbers were giving their lives".

After the war, Rowe first became the Deputy Controller of Research and Development at the Admiralty, and then in 1947 went to Australia as Chairman of the Defence Advisory Committee and Defence Scientific Adviser to the Australian Government, and in 1948 he became Vice-Chancellor of the University of Adelaide. This experience was not a happy one, as Rowe himself related in his book *If*

the Gown Fits (1960). Clearly he found the university a far less ready body to accept his guidance than TRE had been. He thought that universities should be more down-to-earth than they were in meeting the needs of the average rather than the brightest student; but, almost paradoxically, he concluded "It is necessary to replace fearfulness by courage, departmentalism by unity of purpose, egalitarianism by a measure of authority, and to recognise excellence wherever it rears its noble head". And if Australian universities are now better than they were thirty years ago, part of the change may be due to the unpopular stand that he took.

Rowe had never been afraid of expressing an unpopular view, but he felt the unpopularity all the same, when it led to loneliness as in Australia. It was here that he valued especially the support of his wife, Mary, whom he married in 1932: "Far from family and friends" he wrote in *If the Gown Fits,* "she unfalteringly supported me in my years of unaccustomed loneliness".

He came back to Malvern in 1958 and spent some of his later years teaching astronomy to the boys of the College. But—such is the way that a nation sometimes treats its faithful servants—he found it so difficult to live

on his pension that he actually left Malvern again for Malta, where the cost of living was lower. We were glad to see him back again when conditions in Malta proved not to be what he had hoped for. He had the idea of writing a further book, on invention in war, but he found the work too much and handed his papers over to Guy Hartcup, who produced the book as *The Challenge of War* (1970).

When in 1974 the Royal Society organised a Symposium on *The Effects of the Two World Wars on the Development and Organisation of Science in the United Kingdom,* Rowe was unable to attend. The meeting had the atmosphere of a reunion, and it spontaneously sent him a telegram of tribute signed by many of his wartime colleagues headed by the President, Alan Hodgkin, who as a member of TRE himself had made the first flight with airborne centimetric radar in 1941.

His final tribute came on June 18, 1976, when as many of his former colleagues as possible gathered in Malvern Priory for his Memorial Service and to recall his vital contribution as the head of what J. A. Ratcliffe in his Oration fairly described as "one of the most successful research establishments of all time".

R. V. Jones

---

announcements

Appointments

Dr B. J. Mason as President of the Institute of Physics, and **Dr G. H. Stafford** as Vice-President.

Awards

The Royal Society has given the S. G. Brown Award and Medal to **Frank Mackley,** a Director and Chief Engineer of J. T. Mackley & Co. Ltd, for his work on the development of the Hover Platform.
The John Scott Award for 1976 has been made to **Professor Cyril A. Clarke** and **Drs Vincent J. Freda, John G. Gorman** and **William Pollack** for their work on the prevention of Rhesus Haemolytic Disease.

Meetings

October 29, **Hybrids in Botany, Horticulture and Agriculture,** Leicester (Dr C. A. Stace, Botanical Laboratories, The University, Leicester LE1 7RH, UK).
November 15–18, **Weeds,** Brighton, Sussex (Mr W. F. P. Bishop, Frank Bishop (Conference Planners Ltd), 74

Person to Person

A survey is being initiated into the distribution and types of symbiotic *Chlorella* sp. in Britain. Information is needed on precise locations of natural habitats of green Hydra, *Paramecium bursaria, Stentor,* Sponges and other invertebrates definitely containing *Chlorella.* Please write to D. C. Smith, Department of Botany, The University, Bristol BS8 1UG.

Professor and wife wish to find small, furnished flat in S. Kensington or Belgravia, Dec. 15–30, 1976. Will exchange fully furnished 3-bedroomed house in Georgetown, Washington, D.C., Dec. 14—Jan. 6. Address replies and enquiries to D. N. Robinson, 1237 37th Street, N.W., Washington, D.C. 20007.

London Road, Croydon CR0 2TB).
November 24, **Whole Body Counting and Scanning,** London (The General Secretary, British Institute of Radiology, 32 Welbeck Street, London W1M 7PG).
December 12–17, **Automated Cytology,** Pensacola, Florida (Paul F. Mullaney, Chairman, Fifth Engineering Foundation Conference on Automated Cytology, PO Box 208, Los Alamos, New Mexico 87544).
April 14–15, 1977, **Easter Meeting of the Society for the Bibliography of Natural History,** London (Meetings Secretary, Mrs J. A. Diment, Palaeontology Library, British Museum (Natural History), Cromwell Road, London SW7 5BD).
April 18–20, 1977, **Stem Cells and Tissue Homeostasis,** Manchester (Dr R. J. Cole, School of Biological Sciences, University of Sussex, Falmer, Brighton BN1 9QG).
May 9–11, 1977, **Biosystematics in Agriculture,** Beltsville, Maryland (Dr James A. Duke, Publicity Committee, BARC Symposium II, Plant Taxonomy Laboratory, Room 117, Building 001, BARC West, USDA, Beltsville, Maryland 20705).

A place for science in the power game

As a highly respected scientific pressure group, Pugwash might reasonably have expected some publicity for the meeting it organised in London last week to discuss the fast breeder reactor. The advent of fast reactor technology is already becoming acknowledged as perhaps the major social issue of the decade in the developed world. In Britain, the Royal Commission on Environmental Pollution (RCEP) had placed the controversy firmly in the public spotlight with the publication of its report on nuclear power only six days before the Pugwash meeting. But Pugwash went unmentioned by the media. The meeting was ignored by the television corporations, passed unreported on the radio, and received not a single column inch in any of the national daily newspapers. It is worth considering why.

To some extent the unbroken media silence must be attributable to the very timeliness of the meeting, coming as it did so hard on the heels of the RCEP report. There is little press copy to be had from issues already extensively covered less than a week before. Admittedly, the meeting covered most of the contentious ground, and presented both sides of the argument. But it is worth considering whether the media's disregard might reflect more than just a fickle boredom with recurring issues. It could, for instance, say something about the rôle of Pugwash in the widening debate on the fast breeder. The part that it played in educating both scientists and politicians on nuclear matters during the 1960s stands as solid testimony to its influence at that time. But does it wield such influence still?

Pugwash itself apparently believes that it can still reach politicians by going through the back door. Last week's invitation list did not include the UK Energy Secretary, Anthony Wedgwood Benn, or his two political advisers. As for the House of Commons Select Committee on Science and Technology, members were invited, but it looked very much as though most, if not all, of the invitations were treated with disdain. Moreover, of the top personnel from the Department of Energy (DEN) only Walter Marshall could be seen, and his silence and early departure gave no clue as to whether he attended as Chief Scientist at the DEN, Deputy Chairman of the UK Atomic Energy Authority, or simply to assess any new opposition in the nuclear power game. Whatever his motives, the 350 or so scientists who confronted him offered little new beyond the contention that a target date of 1995 for a demonstration commercial fusion reactor in Europe could become realistic with more funding.

But in seeking to understand the lack of media interest not all of the blame can be apportioned to Pugwash and absent politicians. Many of the scientists present would do well to reflect what their true contribution to the fast breeder debate has been. For while it was encouraging to note the number of well-known and respected members of the scientific community eager to discuss all the implication of fast reactor technology, and even prepared to talk seriously in the corridors of the untapped potential of alternative technologies, there remain far too many who see fast breeder reactors simply in terms of load factors, breeding ratios, doubling times, and the like. Consequently, it was left to the eloquent rhetoric of a former Minister of Fuel and Power, Philip Noel-Baker, to set the debate squarely in its political and social perspective, and by the time of his closing speech many of those who ought to have been listening had seen their fill of graphs and tables and dispersed to their buses and trains.

Such a cold approach by the science community makes a mockery of the frequent protests, voiced as loudly as ever during last week's meeting, against media misrepresentation of major science-based issues. For those members of the press whom Pugwash bothered to invite, the choice was simple and clear: they could either search for sensational angles, or settle for reporting a mundane and sometimes technically bewildering meeting exactly as it happened. Evidently, they responsibly chose to do the latter and ran up against their colleagues' more newsworthy competition. Which leaves science with a sad alternative: sensational media coverage (remember the Williams report on genetic engineering?), or no media coverage at all.

And that is the crunch. For though it has wide and inescapable political and social ramifications, the debate over the fast breeder reactor is firmly rooted in science, and it surely falls on the science community to ensure that the public and their politicians are kept intelligibly informed by an impartial, well-briefed press. Sir Brian Flowers, former Chairman of the RCEP and a board member of the UKAEA, recently stressed how the rules of technological advance have changed in the developed world: it is no longer sufficient for each new step forward to be technologically realisable unless it is also socially acceptable. It is in assessing that social acceptability that science must assume a crucial rôle.

Only three weeks ago it looked very much as though the powerful British nuclear lobby might, rightly or wrongly, force an early decision in favour of the UK's first demonstration commercial fast breeder, CFR-1. Now the RCEP, in giving the reactor only qualified approval in its new report, has won a delay of a few months at least for nuclear power critics. Mr Benn and the nuclear lobby have been pushed against the baseline, and the ball has fallen to those in the other court. For the rest of the match it is the scientific community that should act as umpire. And from that position it must ensure that whoever comes out on top, every single blade of grass is covered. □

Small isn't beautiful either

AT the beginning of January 1976, one pound sterling would buy 2.02 US dollars. Last week you could get only $1.66 for a pound; at one stage as little as $1.63. This depreciation of nearly 20% in sterling is making for grim news almost everywhere in Britain; the laboratories of British universities will feel the pinch, for the Science Research Council (SRC) is almost certainly facing an acute financial crisis.

Within the past year the SRC, in common with many organisations dispensing government money, has had to accept cash limits on its activities, that is to say, it has been given a firm sum of money above which it may not commit itself. The SRC has also, of course, been told by the Advisory Board of Research Councils (ABRC) to trim its spending over the next five years by 10% in real terms. To achieve this, the so-called big sciences have been told that many major national projects cannot go ahead, and international facilities such as CERN, Institut Laue-Langevin and the European Space Agency have accordingly increased in importance. Subscriptions to these international organisations amount to about 30% of SRC's budget, and go up every time the pound goes down.

So now, in order to fulfil its international obligations, the SRC is looking for £6 million more than it had budgeted for—with cash limits in operation to prevent it from simply going to the Treasury and asking for the difference. In fact the Treasury is prepared to sympathise to the tune of £2 million, so the SRC is left either to cut £4 million off domestic expenditure, to sweet-talk CERN into letting us pay less, or (the most likely) some combination of these two.

What can be done on the domestic scene? The SRC has relatively few options open, in that much of its money is committed for long-term projects already half-done. Vulnerable spots are therefore new grant proposals and new research studentships. The new crop of research students is barely large enough to satisfy demand; even some candidates with first-class degrees have not yet been placed, and eight hundred candidates with upper-second degrees are going to be disappointed. On the research grants side the scene is even grimmer. At the last round of grant-giving there were significant numbers of applicants who failed even

though their application had been rated in the top category. This had never happened before. One of the options that the SRC is bound to be considering is that the next round of proposals (there are three a year) is simply dropped altogether.

In the somewhat longer run, SRC may have to get down to the serious business of making some of its own staff redundant. It employs nearly 3,000 in six laboratories and headquarters. With generous redundancy terms to meet, the council is unlikely to benefit in the short term from such action, but it might be encouraged by the contents of a forthcoming Confederation of British Industry document 'The Road to Recovery' which, according to *The Economist*, will suggest that industry could create a million new jobs, into which a hundred well-qualified SRC employees could surely fit.

The irony of the SRC's present crisis is that if grant giving is seriously curtailed it will be the 'small sciences' which will suffer, in order to keep the big sciences going—hardly ABRC's intention when it looked to more use of international facilities by Britain's big scientists. □

More on Argentina

CONCERN about the plight of scientists in Argentina has grown steadily since we published a letter from Italian academics on the subject (July 22). It has become increasingly clear that a country whose scientific skills have been widely acknowledged and respected is in the greatest danger of abandoning its intellectual traditions and becoming a scientific backwater. The number of academics and workers in research institutes who have lost their jobs is growing all the time, and some fairly detailed documentation is emerging.

Inevitably, when left- and right-wing extremists are at loggerheads, some of those dismissed or detained will have had a history of political activity, and thus will be vulnerable to charges of using their position for political rather than academic purposes. But many others, relatively apolitical, seem to have been caught in the crossfire. When military politicians such as General Vilas can say that "until we can cleanse the teaching area, and professors are all of Christian thought and ideology we will not achieve the triumph we seek in our struggle against the revolutionary left", there is little hope that anything resembling academic freedom can survive.

Even the Atomic Energy Commission does not seem to have escaped the general purge. At least eight employees are known to have been arrested: Carlos Calle, Dr Domi and his wife Maria Dipace, Pedro Landeyro, Antonio Misetich (who worked for a time at MIT's National Magnet Laboratory), Santiago Morazzo, Pedro Victoria. We know a little about their situation because Dr Virgilo Victoria-Troncoso, brother of Maximo Pedro Victoria and an ophthalmologist at the University of Ghent, Belgium, has recently been to see his brother (who has worked overseas, for example at Lawrence Radiation Laboratory) in prison in Argentina. He reports that the scientists, physicians and psychiatrists he saw in the Villa Devoto prison in Buenos Aires were kept in 35 cells, five to a one-bedded cell with dimensions 3m by 2.40m. They have been without meat for months. They have no legal representation, no charges have been brought against them and it is impossible to bring a writ of *habeas corpus*. "My brother", says Dr Troncosco, "has been beaten up and he has lost teeth. The situation is worse than in a novel of Kafka".

It is difficult to see why nuclear scientists have been imprisoned except

that there was a technical difference of opinion some time ago in which Dr Victoria, amongst others, was involved. It concerned the choice of fuel for a new reactor, and Dr Victoria and colleagues favoured natural uranium whereas a group of military scientists favoured enriched uranium. The former view prevailed and it may be that the military, which has in large part taken over leading civilian positions, is determined to settle an old score.

What can scientists do in the face of the widespread breakdown of all the conventional respect for the intellectual life and academic freedom? Amnesty International (53 Theobalds Rd, London WC1) can provide extensive lists of missing Argentinian academics who could be written to—preferably in Spanish, because the authorities are unlikely to pass on anything in English. Dr Victoria's address, for instance, is now 649, Unidad 2, Sierra Chica, Province of Buenos Aires, Argentina. A letter to Excelentisimo General Jorge Rafael Videla, Presidente de la Republica Argentina, Casa Rosada, Buenos Aires, Argentina couldn't do much harm. And maybe it is going to be necessary to try and find academic places outside Argentina for these victims of extremist politics. □

Watershed for poisons

Colin Norman reports from Washington on US efforts
to legislate on the controversial matter of toxic substances

A FEW years ago, a group of industrial chemicals called polychlorinated biphenyls (PCBs) were enjoying spectacular commercial success, finding their way into a vast assortment of products ranging from electrical transformers to copying paper and food packaging. In fact, PCBs are so ubiquitous, and so indestructible, that virtually everybody living in an industrialised country now has detectable levels of the chemicals in his or her tissue. They are even turning up in human breast milk, they are present in the flesh of fish from many lakes and rivers, and they have been found in the bodies of animals from such remote places as Greenland. A number of studies have shown that, when fed in relatively large amounts to animals, they may cause cancer, reproductive disorders, metabolic abnormalities, hair loss, skin eruptions, and other health problems.

Use of these versatile substances has now been curtailed in the United States. After five years of trying, the US Congress has finally passed landmark legislation requiring many manufactured chemicals to be screened and tested for toxic effects before being marketed. The grim saga of how PCBs became widespread environmental contaminants long before their health hazards were fully known is, unfortunately, far from unique. Other well-known cases include vinyl chloride, asbestos, bis(chloromethyl) ether, carbon tetrachloride and so on. And, though such occurrences are far from new (one novel theory even suggests that the decline of the Roman Empire may have been caused, in part, by lead poisoning from cooking utensils and wine vessels), they seem to be growing more numerous.

It is not surprising that there are such frequent outbreaks of panic about the toxic effects of widely used chemicals. Civilised man is surrounded by thousands of synthetic chemicals, few of which have been screened for toxic effects, and even fewer have been exhaustively tested. According to an estimate published last year by the Manufacturing Chemists' Association, for example, some 6,500 new chemical products reach the market each year. Yet only about 3,000 chemicals have even been tested for cancer-causing properties, in spite of widely accepted estimates that between 60 and 90% of human cancer is caused by environmental factors.

In short, until now, chemicals have generally been accorded the same legal rights as people—they are assumed innocent until proven guilty. But last week, after years of argument, Congress took a significant step toward writing a new bill of rights for manufactured chemicals. It passed legislation, known as the Toxic Substances Control Act, which gives the Environmental Protection Agency (EPA) broad authority to require pre-market testing of some compounds, and the power to seek a court order to keep potentially hazardous chemicals off the market.

The bill is a landmark piece of legislation for which environmentalists, some trades unions and health scientists have long fought. For the first time, it gives the federal government the power, in theory, to keep an eye on all manufactured chemicals, and it provides at least a coarse filter to screen out potentially troublesome compounds before they do much damage. In the colourful words of Senator Warren Magnuson, a key supporter of the bill, it will "no longer allow the public or the environment to be used as a testing ground for the safety of these products".

Five years of effort

It took five years of intense argument in Congressional committee rooms and behind the scenes to get any toxic substances legislation passed at all. The bill's genesis was a 1971 report by the Council on Environmental Quality, which noted that although some classes of chemicals (such as pesticides, drugs and food additives) are regulated by individual laws, there is no federal authority to control the thousands of other compounds which flood the market each year. A year later, the Senate and the House both passed toxic substances bills, but they were markedly different, and intense lobbying by industry and by the Nixon Administration prevented a final version being passed. The same thing happened in 1973 and 1974, and until very recently it looked as though Congress would again fail to reach agreement on the legislation. The deadlock was broken during what one participant described last week as "three intense bargaining sessions" between Senators and House members.

What emerged was consequently very much a compromise measure, a fact which is reflected in comments last week from spokesmen for groups on both sides of the fence. "We support the bill as a workable compromise", says Linda Billings, a Sierra Club lobbyist who has been following the legislation throughout its tortuous journey through Congress. She noted, however, that it contains a number of weaknesses and loopholes which will keep many corporate lawyers busy. Similarly, a spokesman for the Manufacturing Chemists' Association (MCA), the industry's lobbying arm, said that the bill, "while tough", is acceptable to the industry. He added, though, "I wouldn't say that we are real happy with everything that's in the bill". Spokesmen for some trades union organisations have also expressed satisfaction that a bill has finally been passed.

One lingering doubt, however, is whether or not President Ford finds the bill to his liking. Throughout the Congressional fight on the legislation, the Administration has lobbied against the measure, arguing that it would be expensive and would constitute too much government regulation of private industry (a theme frequently sounded in Ford's campaign rhetoric). But, since the bill now has broad support, and since a Presidential veto would give Jimmy Carter a gilt-edged campaign issue to exploit, Ford is expected to sign the legislation.

A key feature of the bill—which Ms Billings describes as a "watershed"—is a requirement that EPA must be informed at least 90 days before any new chemical compound is placed on the market, or an existing compound is sold for a new use. (The bill specifically exempts chemicals produced in small amounts for research purposes, it should be noted.) The manufacturer must also send along whatever information he has on the toxicity of the compound, the amount to be manufactured, the likely human exposure, and so on.

Then, if the EPA Administrator decides that the compound is likely to pose a hazard to the environment or to human health, or if he finds that there is insufficient information to judge the hazards, he can issue an order restricting or banning sale of the compound, at least until the required test data is available.

Sticking point

The EPA Administrator is clearly given very flexible authority to determine which compounds should be exhaustively tested, and which should be let through the filter. That aspect of the bill has been the chief sticking point which has held up final agreement in Congress for the past five years. The deadlock was broken by writing into the bill a provision requiring that the Administrator's decisions

must be reviewed by the courts if a manufacturer feels he has been unfairly treated. In short, if a manufacturer objects to an EPA ruling, the agency must seek a court injunction to put the ruling into effect. If EPA doesn't go to court, the ruling would be nullified.

Although that provision may seem like a huge loophole through which smart corporate lawyers can emasculate the legislation, many observers feel that in fact it will be relatively easy for the EPA to obtain an injunction. All that would be required, according to the bill, is a showing that the compound "may present an unreasonable risk of injury to health or the environment", or that it "may reasonably be anticipated to enter the environment in substantial quantities", or that there has been insufficient testing for the hazards to be "reasonably determined or predicted".

Thus, the wording is so general that EPA should have a fairly easy time in proving its case, a fact which made the provision acceptable to environmentalists. "We don't like it", Ms Billings said last week, "but we recognise that it is the best we could get". Similarly, the chemical industry sees some merit in the provision. James Hanes, chief counsel for Dow Chemicals, who has been one of the most outspoken opponents of the bill, told *Nature* that the idea "at least gives the industry the chance to contest the decision out in the open", though he noted that "I don't really see the courts requiring an overwhelming burden of proof here".

If EPA obtains its injunction, it can then specify what tests it requires, and the industry would then be in the position of having to prove that a product is safe before placing it on the market.

Priorities committee
As far as existing chemicals are concerned, the bill requires the setting up of an inter-agency committee charged with the task of drawing up a list of chemicals whose toxicity is open to question. The committee will assign priorities to chemicals and, within 12 months, the EPA Administrator is required either to issue orders for testing the top 50 compounds, or to explain why he feels testing isn't necessary.

The bill will clearly place a huge burden on EPA, which is required to go through masses of data on thousands of chemicals, decide which need testing and which should be allowed on the market, go to court to obtain injunctions, and exercise considerable judgment on what actions are necessary. Therein lies the bill's greatest potential weakness, according to many observers.

To carry out this Act, the bill authorises expenditures of only $10 million this year, rising to $12.6 million next year and $16 million the year after. Those figures should be compared with the $125 million which EPA was budgeted to enforce the Clean Air Act, or the $200 million a year budget of the Food and Drug Administration. It is difficult to see how EPA can effectively enforce the Toxic Substances Act with such small resources, and some observers have suggested that the upshot will be that EPA will be forced to concentrate on a few chemicals and allow many to slip through the net.

Monetary matters have also been a major source of concern to the chemical industry throughout the long Congressional fight over the bill, though for different reasons. There have been numerous studies of what the bill may cost the industry in terms of testing facilities, legal costs and administrative requirements. Large-scale animal tests are very expensive to conduct—according to estimates given in testimony before a House committee by officials from DuPont, for example, a complete battery of tests can cost up to $500,0000. Such figures have thus led to arguments that the bill will stifle the development of important chemicals which may only have a short production run. Some of that concern has, however, been reduced by the provision exempting research chemicals from the bill.

Assessing the cost
As for the total impact of the bill on the industry, Dow Chemicals has so far come up with the highest figure, estimating that the legislation will cost chemical manufacturers up to $2,000 million a year. The MCA thought that the cost would probably lie between $360 million and $1,300 million, while EPA suggested that it would only be about $80–140 million.

Whatever the cost may turn out to be, it will be relatively small compared with the massive total sales volume of the industry, or with the costs of treating cancers and other diseases related to environmental factors. The large cost of animal tests, moreover, should provide a strong incentive for industry to develop accurate, short-term *in vitro* tests, such as the system developed by Dr Bruce Ames at the University of California. Already, many large firms are looking closely at such systems, and are using them as rapid screening devices, essentially to develop priorities for animal tests.

It now remains to be seen whether or not President Ford will sign the bill. If he refuses, he not only runs the risk of the matter developing into a campaign issue, he would also probably do the chemical industry more harm than good. This bill swept through the Congress with support from both Democrats and Republicans, there is certainly strong public pressure for such legislation—especially since the panics over PCBs, vinyl chloride and so on—and the polls now indicate that the Democratic majority in the Congress will be much stronger next year. All those considerations suggest that an even tougher bill would emerge from the next Congress if this bill expires. That is one reason why the chemical industry has decided to give this bill at least its lukewarm support.   □

USA

Science Court on guard

A number of prominent scientists and government officials met last month to discuss a proposal to establish a 'Science Court' in the United States, to examine complex scientific issues which have a bearing on public policy. The proposal (see box) has recently been receiving considerable publicity in the United States. Wil Lepkowski reports from Leesburg, Virginia

AT first blush, the idea of a Science Court sounds a bit intriguing if not a mite exciting: a forum of distinguished scientists and engineers gathered together in court for sifting, filtering, and distilling contradictory facts into one sparkling supernatant of truth. The world of political decision-making is so messy, proponents of the science court assert, that officials would welcome a little plain, unvarnished truth from a forum that has no political stake in public issues.

This is a fairly accurate, if not precise, description of the idea behind the science court, which is mainly the invention of Dr Arthur Kantrowitz, president of the Avco Research Laboratories. For about a decade, Dr Kantrowitz has been tirelessly giving on and off the cuff speeches on the need to establish such a court and, after a good deal of effort, he finally succeeded in securing government support for a colloquim on the subject. It was held last month at the Xerox Cor-

poration Training Center in Leesburg, Virginia.

About 300 people came, representing law, science, engineering, industrial management, government administration, social science, and general subjects. No one from the humanities was observable, unless one counts anthropologist-oracle Margaret Mead. Dr Mead was the sole bearer of mirth into what was otherwise a dull, juridical gathering. She was in general kind, affirmative, but at one point called the affair a publicity meeting. That may be unkind to Dr Kantrowitz, who in sincerity is outranked by no one. But the whisperings around Washington in that small circle called the science policy community (not to be confused with the working scientists) is to give Kantrowitz his day and let the whole idea collapse under its own unworkable pretensions.

Nothing was really decided at the colloquium. Individuals theorised how it might work, what might be the snags, how the right panels of experts would be picked according to the issue under debate, who would choose the experts, how information would pass between the court and the regulatory agency subscribing to it, what the procedures would be for cross examination, how the public would participate, whether the system already works well enough. In short, the meeting was a mélange of speculation. It is probably accurate to say that Dr Kantrowitz, who heads a science court task force set up by a now-disbanded White House science advisory committee, will get a little money to do a science court experiment with a regulatory agency. That's what he intends to do. In fact, the colloquium's point was to hash over the elements of such an experiment.

Two concrete issues were discussed around which an experiment might be devised: food additives and nuclear power. Both issues have enjoyed and lamented over huge outpourings of fact, opinion, and acrimony over many years. In both cases the matter boiled down to the free flow of information —facts withheld by government and industry that no science court could pry loose. And anyway, one could make a fairly substantial case for the "fact" that science courts already exist— through panels serving regulatory agencies, through the National Academy of Sciences, through existing legal courts, in other words through a system within the American democracy that, after all is said and done, does indeed spot the scallawags.

It is certain that the experiment will be tried. The President's science adviser, Guyford Stever, gave it cautious endorsement. Even the Commerce Secretary, Elliot Richardson, said it might as well be tried, although

Science court: specifications

A top-level science advisory committee, established by President Ford last year to prepare the way for the re-establishment of a White House science policy office, recommended that the Science Court idea should be given serious consideration. It established a task force, chaired by Dr Arthur Kantrowitz, which has drawn up the following specifications for the court:

Initiation

1. The experiment will begin with agreements between the Science Court Administration (SCA), the host institution and the funding institution concerning the plans for the experiment.
2. The first step will consist of contacts with regulatory agencies and others to search for a suitable issue. The first issue will be presumed to be bi-polar and will be presumed to be a value laden decision which must be made by the agency and in which the scientific facts are apparently in doubt.
3. An agreement with the regulatory agency to supply necessary legal powers for uncovering non-public information if necessary.

Organisation

4. The SCA, equipped with an issue and funds, will seek Case Managers (CMs) for each side of the issue. The work of the CMs will be funded by the SCA.
5. The SCA will issue procedures, suggested referees, and a panel of prospective judges.
6. The CMs will review these suggestions. The experiment will go forward only after both CMs agree on the referee, the judges and the procedures. A signed agreement to proceed will commit the SCA, the funding agency, the regulatory agency, the CMs, the judges, and the referee.

Suggested procedures

7. The CMs formulate a series of factual statements which they regard as most important to their cases. Such statements must be results or anticipated results of experiments or observations of nature.

The statements should be ranked in order of importance assigned by the CM.
8. The judges examine each statement to determine that it is a relevant scientific fact.
9. The CMs then exchange statements. Each side is invited to accept or challenge each of the opponents statements.
10. The list of statements accepted by both sides will constitute the first output from the SC.
11. Challenged statements are first dealt with by a mediation procedure in which attempts are made to narrow the area of disagreement or to negotiate a revised statement of fact that both CMs can accept.
12. The mediated statements are added to the SC's output. Those statements which remain challenged are then subjected to an adversary procedure.
13. CMs prepare substantiation papers on statements remaining challenged and transmit these to the judges and the opposing CMs starting with the first (most important) challenged statement.
14. The substantiation is cross-examined by opposing CMs and judges and contrary evidence is presented and cross-examined.
15. A second attempt to negotiate a mediated statement is made and if successful this statement is added to the Science Court's output.
16. If this is not successful the judges write their opinions on the contested statement of fact.
17. This procedure is repeated for each of the challenged statements.
18. The accepted statements plus the judges' statements constitute the final output of the procedure.

it shouldn't be confused with a court of law. Stever did imply, however, that one shouldn't expect too much.

The mythology behind the science court concept does deserve an analysis that the colloquium did not consider. The question concerns reasons why the science court idea is deemed so urgent as to have a whole colloquium called to discuss it. Industrialists seem to want it because they are convinced that strident opposition to technology by environmentalists are stifling their own freedom to make fair profits. Scientists and engineers seem to be for it because it gives them the chance to establish order in forums often run emotionally. Environmentalists and public interest figures are against it because they fear their positions will be muted by a scientific elite that has little feeling for human needs. And lawyers are against it because they feel court procedures, if done right, hear all sides to a case, within human limits of patience, tolerance, and intelligence.

On the scope of reality, the science court seems of little importance beyond a possibly interesting intel-

lectual exercise. Regulatory agencies make literally dozens of decisions daily. Hundreds of scientific panels are called to aid in those decisions—little courts, if you will. The National Academy of Sciences runs a series of Science Forums every few weeks or so in which the public actually does debate with scientists the validity of scientific findings relevant to public issues. The administrators of the Science Forum aren't pretentious enough to believe their procedures heavily influence decisions. But some follow-up studies do indicate that the debates raise the personal and collective consciousness to a degree satisfactory to anyone wise enough to understand that people aren't swayed very easily.

The trouble is that no one quite knows what to think of the Science Court. It is a difficult target to attack because the subject is virtuous. Many simply don't feel it is worth much. One might call the feeling the accumulated wisdom of those who have been through the course of using scientific facts in the pursuit of the common weal.　　　　　　□

SWEDEN

Swept in on the green wave?

From Stockholm, Wendy Barnaby assesses what the result of the recent Swedish general election could mean for science in that country

AFTER 44 years in power, Sweden's Social Democrats have been defeated at the polls. Their general approach to science was not discussed in the election campaign. But scientific assessments became very important in the debate which dominated its final stages: whether or not to push ahead with the Social Democrat's plan to make Sweden the world's largest producer of nuclear energy *per capita* by 1990. The scientific input was dramatised by a telegram sent two days before the election by the American Union of Concerned Scientists to the Prime Minister, Olof Palme, challenging his view that nuclear waste disposal problems are on the way to being solved.

Although the non-socialist parties which will now form a coalition government are split on the nuclear issue, the pre-election battle lines were drawn between the Social Democrats on the one hand and the anti-nuclear energy Centre Party, the largest non-socialist party, on the other. And in spite of the enormous organisational and governmental machinery at the Social Democrats' disposal, it was the arguments of the non-establishment scientific community which ultimately appealed to the Swedish voters.

The Centre Party, traditionally the political voice of the Swedish farmer, has in the past few years been the champion of the so-called 'green wave': demands for decentralisation of population and power, environmental protection, the careful use of finite resources and—in the energy field—the development of reliance on renewable resources such as wind and solar power. According to the party's spokesman on scientific affairs, Bengt Sjönell, one of the new government's aims will be to bring the Swedes into balance with nature as far as energy resources are concerned. The first test of how this rather idealistic hope is to be realised will come less than 2 months after the formation of the new government, early in October, when the squat Barsebäck 2 nuclear power reactor, becomes ready for fuel insertion.

The Centre Party's leader (the new Prime Minister), Thorbiörn Fälldin, is personally and publicly committed to stopping the loading of fuel into any more reactors. (He has also promised close down those five reactors already in operation.) Nobody doubts

the strength of his commitment. But the problem with Barsebäck 2 is how fuel loading can be prevented. One proposal is to prohibit the insertion of fuel by passing a succession of laws each valid for one month at a time until compensation can be negotiated with the reactor's owners, who have demanded $445 million from the government. The trouble with this is that the reactor is owned by a private company, Sydkraft, and for a non-socialist government to make its legislative debut with such regulation of private industry would be politically unpalatable. To get around this difficulty, it has been suggested the Nuclear Power Inspectorate—a state body which checks reactors under construction and in operation to see that they conform to agreed standards—could prohibit fuel loading at Barsebäck 2. But that organisation is itself not at all happy with the idea. Its brief covers purely technical areas, and although the checklist for the reactor has not yet been entirely worked through, there have so far been no technical hitches. Dr Arne Hedgran, a representative of the inspectorate recently made it quite clear that the organisation does not want to become involved in the government's political problems.

The tangle over Barsebäck 2 is the most pressing problem faced by the new government; not only because of the time factor but also because the Centre Party's credibility will be greatly affected by its handling of the situation. In its very first days of power it will have to come to grips with the issue held to be largely responsible for its victory.

The influence of the already-powerful environmentalist groups will no doubt increase under the new government, partly because of official willingness to listen, but also because of significant gains by the environmentalists in the local elections held at the same time as the national ones. One such local result has particularly far-reaching national and international implications. A victory by groups opposed to the mining of uranium near Billingen, 350 km south-west of Stockholm, means that uranium reserves estimated to be the West's largest in the next economical price range cannot now be exploited for at least three years. The mining firm LKAB had been hoping to mine them either for domestic use or for export.

Beyond the immediate problems, what sort of initiatives in science policy can be expected from the Centre Party? It should be mentioned that Sweden

Fälldin, committed

has no ministry of science: the administration of science instead falls to various other ministries. Those most concerned are Education, Industry and Agriculture, but even they exert little pressure on the direction of research. Once funds have been allocated to the extraordinarily varied array of institutes and research bodies, each research group is practically autonomous in deciding how they are to be spent. Consequently, it is often hard to see whether Sweden has any science policy as such at all.

One trend which is, however, apparent in all this diversity is the priority given to applied research over basic research. This emphasis has been responsible for the present plight of so-called basic technical research, which has been practically squeezed out of existence in Sweden. Bengt Sjönell recently declared that the Centre Party wants to strengthen the relative position of basic research, and stressed specifically that support for basic technical research would be increased. But the party's intentions have not yet been expressed in detailed proposals, or even in concrete plans, though in tertiary scientific education, the party will try to tempt students back to research in the natural sciences, whose current unpopularity has been raising questions about Sweden's long-term scientific research capabilities. But Sjönell seems unaware of the potential of the International Energy Institute due to begin work here next year. Although the government will have no control over the work programme at the institute, it is a reasonable bet that much of the research done there will be directly relevant to official aims. And for a government whose interest in science is concentrated in the energy field, the value of such research should not be underrated. In other areas of Swedish science, the prognosis is business as usual. ☐

IN BRIEF

NIH budget veto overridden

A few days before it adjourned for the November elections, the US Congress firmly overrode a Presidential veto of a budget bill for the National Institutes of Health (NIH), thereby increasing NIH's biomedical research funds by more than $200 million in the fiscal year which began on October 1. The bulk of the increase will go to the National Cancer Institute (NCI), whose budget will soar from $762 million to $815 million. Also set for a large increase is the National Heart and Lung Institute, which will get $397 million, up from $370 million last year. The final budget levels represent a considerable victory for supporters of the cancer programme, for the Ford Administration had recommended that the NCI's budget be held approximately constant and that the other NIH institutes be given modest increases.

Air Act problem

The US Congress has handed the automobile industry a tangled legal problem by failing to pass amendments to the Clean Air Act before it adjourned last week. The amendments would have relaxed emission control standards for 1978 model cars, which the industry is already gearing up to produce. Members of Congress failed to agree on the bill, and it died when the session ended in the early hours of Saturday morning. Thus, unless the new Congress approved laxer standards when it convenes next January, the industry will be faced with the choice of producing illegal cars or closing down. That, in fact, is exactly the situation that the industry has been trying to engineer. It believes that the threat of wholesale closures and high unemployment in Detroit will be sufficient to force the new Congress to meet its various demands.

UK changes

Under the rearranged provisions for science policy formulation in Britain, the UK government has appointed a 37-year-old biochemist from Essex University, Professor John Ashworth, to the Cabinet Office's Central Policy Review Staff for a two-year period.

Elsewhere, Professor J. L. Gowans, 52, will be the next Secretary of the UK Medical Research Council, taking over from Sir John Gray in April 1977. He is presently Director of the MRC's Cellular Immunology Unit.

ALTHOUGH the sixth report of the Royal Commission on Environmental Pollution, and various informed comments on it, have shown that there is widespread concern in Britain about the rapid development of nuclear power in general, and of the fast breeder reactor in particular, those responsible for the safety of atomic energy and nuclear fuels have come well out of the controversy. Engineers, scientists and those working in nuclear power stations have behaved responsibly in this respect, and have scrupulously observed the most rigorous safety precautions. Many of us may doubt whether we need this massive development in electrical generation which the Department of Energy thinks necessary, when we already have 50% more capacity than the peak load, and when demand is falling, but that is another matter. There is little doubt that if the nuclear generating industry is expanded as planned (*sic*) the same high standards of care will continue to be exercised, though the statistical chances of accidents must inevitably increase.

Therefore if we are to have more nuclear power stations, we must constantly review and improve safety precautions. At present there is a series of safety regulations laid down and agreed internationally; unfortunately some of these are not entirely practical. Thus the specifications for the containers, each weighing some 45 tonnes, in which nuclear fuel is shipped around the country, are intended to ensure that the risks of an escape of radiation or radioactive materials are very small indeed. The vessels are designed to resist such mishaps as a fall of several metres, though there is doubt whether they are strong enough to remain intact if the train carrying them was derailed at a high speed.

The regulations state that they must be able to resist a temperature of 800 °C, uniformly applied to the

---

Protection practice
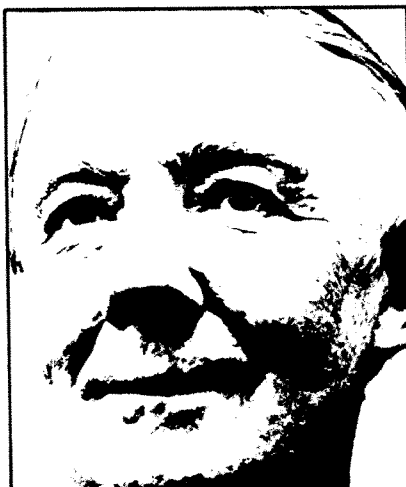
KENNETH MELLANBY

---

whole of their surface. Unfortunately the experts of British Nuclear Fuels Ltd have found it impossible to devise a technique by which the containers can be subjected to the conditions laid down by the international authorities. They have spent—or wasted—some half a million pounds trying to devise a method of giving exactly the correct exposure to this high temperature. As this sort of temperature exposure appears to be impossible to produce, it is difficult to understand why this particular standard was set. We wish to be sure that our nuclear fuel is safe when the containers are exposed to the sort of hazard which, in practice, is likely to be encountered. Uniform conditions of 800 °C continued for twenty minutes do not constitute one of these. It would surely be better to devise a "standard fire" similar to one likely to be experienced by the nuclear fuel when in transit, and to put the containers into it.

A similar attitude to that which sets up these unreal safety standards may sometimes be demonstrated by those responsible for our defence. Some years ago conservationists were urging the War Office to give up an area of Britain's coastline, where the sand dunes were of great natural history interest. The War Office was reluctant to part with the land, giving as an excuse that it was of the greatest importance for military training because it was "unique" and irreplaceable by any other area which was available. When I heard this I could not help wondering how it was to be used for training our forces. If it was indeed unique, then any military exercise performed there would have little relevance to those done on other, quite different, terrains. In fact the only value of this special piece of coastline would be in order to train our forces to defend it. Surely training would be better done on a more typical area, with conditions similar to those found in many parts of the globe?

We need to be protected from danger by safety standards for hazardous industrial processes, and by the possession of efficient military defences. In both cases we should first decide exactly what we need to be protected from, and then we can set up practicable standards to which we can work.

news and views

Doping amorphous silicon

from Andrew Holmes-Siedle

Recently (see *News and Views,* **260,** 667; 1976) a group at the University of Dundee (Spear *et al., Appl. Phys. Lett.,* **28,** 105; 1976) reported that, for the first time, a p–n structure had been made from amorphous silicon only. This achievement demonstrates that, contrary to common belief, amorphous material can be "doped" by impurities. Such a finding effectively eliminates one of the hypotheses which has been aired to explain the suppression of doping in amorphous materials—namely that the valence bond of the dopant atom (which cannot be satisfied within a constrained crystal structure) can usually be accommodated and satisfied by an atom of the disordered matrix. Here, however, is a truly amorphous silicon matrix which can be made either p–type or n–type by the addition of boron or phosphorus, although more dopant is needed than for the case of crystals.

Spear and LeComber have now published some more basic investigations on conductivity and field effects in a much simpler structure, namely a silcon film of only one conduction type (Spear and LeComber, *Phil. Mag.,* **33,** (6) 935-949; 1976) and their new account shows that my earlier comments on the "apparent inefficiency of doping" and my implication that disorder in these silicon networks was still largely "negating" the electrical activity of the dopant atoms was an oversimplification. In fact, quite a large fraction of the dopant atoms is ionised, as we shall see below.

Spear and LeComber deposit the silicon films by glow-discharge decomposition of silane on glass, using phosphine or diborane as dopant gases. This is a well-known technique but is the first intensive study of the product. They found that the conductivity of the films could be controlled in unprecedented fashion for amorphous materials, from 10^{-7} to 10^{-2} (ohm cm)$^{-1}$.

In the past, sputtered films have usually had very low conductivity and any changes achieved (for example by heat treatment—see stippled area of figure) were caused by changes in the density of intrinsic defects, sometimes called "dangling bonds", rather than by true doping effects. Such ability to control conductivity could open up the technological uses of thin-film amorphous silicon—a very desirable thing because of the cheapness and large possible areas of such films. However, there are some formidable, fundamental and technological barriers to be surmounted before devices competitive with single-crystal silicon forms can be created. The efficiency of doping as it affects conductivity is still quite poor, for some fundamental reasons. First, the mobility of carriers in the disordered matrix is lower than for single-crystal semiconductors. Second, many carriers are immobilised in a class of energy states special to amorphous semiconductors ("mid-gap states"). The achievement of Spear and LeComber is to reduce these states until at least a small proportion of the electrons and holes, freed from the dopant atoms, appear in conduction states. Nevertheless, the poor relative efficiency of dopant atoms in producing conductivity is clearly shown up in the figure, adapted from the new measurments of Spear and LeComber (crosses) and other data (full lines, dotted lines). N_D^+ was estimated tentatively by Spear and LeComber from the temperature dependence of conductivity in the amorphous films, with supporting evidence from chemical analysis and photoconductivity measurements. The shaded area shows the difference between amorphous and single-crystal material. While further work may bring these curves somewhat closer together, there is no immediate clue as to how this may be done. In the new material, the density of electron states

A FOOTNOTE to the work reported here is now required because, as might have been expected, the possible use of amorphous silicon films to make photovoltaic p-n diodes has not been ignored. The first to publish is RCA Corporation; others will probably follow shortly. The RCA workers, D. E. Carlson and C. R. Wronski (*Appl. Phys. Lett.,* **28,** 671; 1976), have made thin-film, back-surface amorphous-silicon solar cells on glass by a deposition process essentially the same as Spear and co-workers at Dundee (glow-discharge deposition of silane gas) but the RCA films used are thicker than the Dundee films and have an undoped ("intrinsic") region of thickness about 1.5 μm between the p and n-type regions. The 3.5 cm^2 cells produced have a terrestrial solar conversion efficiency of 2.4% and the authors estimate that the theoretical efficiency of their p-i-n structure should be 14–15%. Thus, the unexpected finding that useful junction barriers can be rather easily produced in amorphous silicon may lead to a surge of effort to produce low-cost, large-area solar energy converters by this method. The economic advantages of thin-film devices over ingot-grown semiconductor devices have been stressed and we may now be avoid the problem of the polycrystalline thin-film approach.

Though surprising advances have recently been reported in fabrication of polycrystalline silicon cells (5% efficiency claimed for large-area polycrystalline films deposited on graphite, *Electronics,* 33, June 24, 1976) the grain boundaries, which constitute a completely unwanted barrier to carrier transport in such films, are dispensed with in amorphous layers.

in the middle of the gap may be as low as 10^{17} cm^{-3} eV^{-1}. This is a huge improvement over previous films, which had densities as high as 10^{20} cm^{-3} eV^{-1} at mid gap, but how much further can we go towards the case of single-crystal silicon, with a density of virtually zero? Some opinion says that the gap in the figure probably cannot be closed much more, but it is still a fascinating question.

Many other fascinating questions will doubtless be raised during further basic physical studies on the new material. With diodes, for example, recombination in amorphous networks can be explored, recombination being the property which most affects gain in transistors. With the doping techniques, we can refine our models of electron and hole transport in regimes not possible before and thereby redress our present ignorance about the environment of substitutional impurities in disordered networks. Out of such work could grow the expertise necessary to exploit this second generation of amorphous films.

Comparison of conductivity against doping curves for single crystal silicon and amorphous silicon (a-Si).
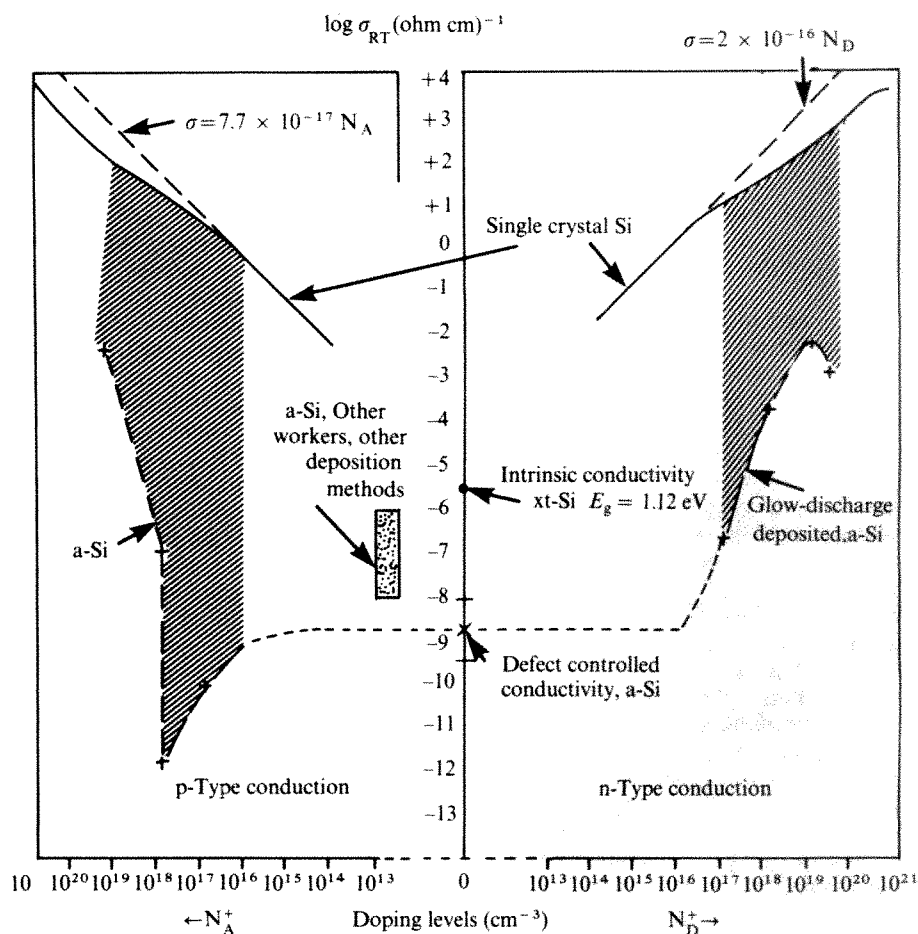
Chlorophyll dimer and triplet states – key roles in photosynthesis?

from Godfrey Beddard

It has been known for some years that the initial photochemical act in photosynthesis is the formation of an oxidant P$^+$, and a reductant X$^-$ resulting from excitation of the reaction centre PX, either directly by light absorption or indirectly by energy transfer from accessory pigments. This reaction centre is composed of a special chlorophyll complex, the electron donor, closely associated with an electron acceptor.

The nature of the reaction centres has been the subject of much interest and speculation and recently experiments using optical spectroscopy and magnetic resonance techniques have resulted in a detailed picture of the bacterial reaction centre P$_{870}$ and the characterisation of the two reaction centres P$_{700}$ and P$_{680}$ of green plants. Chlorophyll dimers have been implicated in all three types and recently electron spin resonance (ESR) spectra exhibiting the CIDEP phenomenon suggest that chlorophyll triplet states are also involved, as described McIntosh and Bolton (*Nature*, **263**, 443; 1976).

Picosecond spectroscopy on bacterial reaction centres has revealed transient absorption changes, implicating both bacteriochlorophyll (BChl), and macteriopheophytin (BPh). The transient forms in 8 ps and decays with a 150 to 250 ps lifetime concomitant with P$_{870}^+$ X$^-$ formation. (Rockley, *et al. Proc natn. Acad. Sci. U.S.A.*, **72**, 2251; 1975: Kaufmann, *et al. Science*, **188**, 1301; 1975). Fajer *et al.* (*Proc. natn. Acad. Sci. U.S.A.*, **72**, 4956; 1976) proposed that the primary electron acceptor was not an iron–ubiquinone complex as was previously thought, but was instead bacteriopheophytin. In an elegant experiment Fajer *et al.* reproduced the transient absorption spectrum of the reaction centre from the difference spectrum; (BChl)$_2^+$ − (BChl)$_2$ plus BPh$^-$ − Bph. Agreement was found for all absorption bands from 350 to 1,250 nm as a transient absorption of (BChl)$_2^+$ has been detected at 1,250 nm. (Dutton *et al., FEBS Lett.*, **60**, 275; 1975).

Triplet state zero-field parameters have enabled Clark *et al.* (*Biophys. Biochem. Res. Commun.*, **71**, 671; 1976) to determine that the two BChl molecules *in vivo* are tilted at 48° and rotated by 78° to one another. Further-

more ENDOR linewidths indicate dimeric BChl species are present, (Feher, *et al., Ann. N.Y. Acad. Sci.*, **244**, 239; 1975; Norris, *et al., ibid.*, 239; 1975) and in addition electron spin polarisation indicates that the primary act is a radical pair formation. (Thurnauer, *et al., Proc. natn. Acad Sci. U.S.A.*, **72**, 3270; 1975). These experiments support a model of the bacterial reaction (in *Rhodopseudomonas sphaeroides*) in which the charge separation process forms the ion pair (BChl)$_2^+$ BPh$^-$ before the reduction of the ubiquinone "primary" acceptor.

The chlorophyll-*a* species of photosystem I reaction centres—P$_{700}$—has a bleachable absorbance at 703 nm and a characteristic ESR signal. *In vivo* ENDOR linewidths indicate that P$_{700}$ contains a (Chl)$_2$ species (Feher, *et al., op. cit.*; Norris, *et al., op. cit.*), and Shipman has recently proposed that this dimer has the form (Chl–ROH)$_2$; with R = H, ethyl or protein. Shipman, *et al., Proc. natn. Acad. Sci. U.S.A.*, **73**, 1791; 1976). This symmetrical dimer has absorbance and ESR spectra similar to P$_{700}^+$. The oxygen in ROH bonds to the Mg of Chl and the

hydrogen bonds to ring V carbonyl of the other Chl. As ROH, RNH, RSH compounds also form dimers this presents the intriguing idea that the *in vivo* dimer may be bound to similar groups in a protein but at present no experiments have been reported to test this idea.

Perhaps the most interesting observation on P_{700} is a transient electron proton resonance (EPR) emission spectrum produced by a single light flash on chloroplasts at room temperature. (Blankenship, *et al., Proc. natn. Acad. Sci. U.S.A.*, **72**, 4943; 1975). The EPR signal is thought to be that of the primary acceptor as it does not correspond to P_{700}^+ or a ferredoxin.

The emission is caused by the two spin states, normally thermally populated to 0.1% of one another, obtaining a significant population difference as a result of a particular chemical reaction favouring one state over the other. This phenomenon is called chemically induced dynamic electron polarisation (CIDEP). Immediately after the radical anion of the acceptor is formed the spin states still have the same populations as the anion's precursor and a CIDEP signal is observed until spin lattice relaxation equalises the spin populations. A triplet state of Chl is thought to be responsible for the emission from photosystem I, as a radical pair, which can also generate CIDEP are unlikely to diffuse together in the reaction centre.

There is still much controversy about the identity of the primary acceptor for P_{700}. Malkin *et al.* (*Biochim. biophys. Acta*, **430**, 399; 1976) maintain that the acceptor is a ferredoxin, but McIntosh and Bolton (*Biochim. biophys. Acta.*, **430**, 555; 1976) and Evans *et al.*, (*Nature*, **256**, 668; 1975) dispute this and claim that a species with ESR signal $g=1.76$ is the primary acceptor and that ferredoxin has a secondary role. The relationship of the signal seen by Blankenship to these acceptor signals is not known. The evidence is not as complete as with the bacterial reaction centre but the initial charge separation in photosystem I appears to produce a $(Chl)_2^+$ species and an unidentified radical anion before the ferredoxin is reduced.

Much less is known of photosystem II reaction centres, P_{680}, than of P_{700}, although there is good evidence that P_{680} is the primary donor, possibly a chlorophyll dimer (Pulles, *et al., Biochim. biophys. Acta*, **440**, 98; 1976). The observation, by McIntosh and Bolton, of a CIDEP emission from Chl in photosystem II is thus of particular interest as this signal can only arise from a Chl triplet state at the temperatures used. They suggest that the triplet state forms very rapidly from the singlet by the formation of a Chl^+

Chl^- species. The primary acceptor would then have to be adjacent to the dimer in order to accept the electron at a rate much faster than the charge recombination for efficient charge separation to occur.

Many similarities between reaction centres are now emerging. All involve chlorophyll dimers; at least two, and perhaps all, contain short lived "transient" acceptors between the dimer and the "primary" acceptor which is perhaps a device to prevent rapid charge recombination of the ion pair. Although both P_{700} and P_{680} appear to involve triplet states formed, perhaps very quickly, after excitation of the dimer species the reasons for preferring triplets over singlets are unknown. Consequently picosecond spectroscopy will be invaluable in observing transients produced in the initial processes and our present knowledge of reaction centres should provide impetus to produce model systems to mimic photosynthetic processes. □

Homostable operons?

from E. G. Richards

READERS who glance through the pages of the more theoretical journals will be aware of the immense activity over the past 15 years on the statistical mechanical basis of the melting of DNA. Certainly every well-informed undergraduate will know that DNA undergoes a cooperative helix coil transition when heated that gives rise to a sharp melting curve. However theory suggests that things are not quite that simple. For instance, a hypothetical DNA molecule consisting of a region of 1,000 A–T base pairs followed by a similar one of G–C pairs would give a melting curve consisting of two sharp steps. To produce a sharp step, theory suggests, requires a 'homostable' region of 500 or more base pairs with no sequences of more than a few adjacent A–T or adjacent G–C base pairs. A DNA molecule consisting of several such homostable regions would then melt in a series of sharp steps, and the melting temperature of each step should reflect the base composition of the corresponding region.

Another and more informative way of depicting melting behaviour is by the differentiated melting curve in which sharp steps now appear as sharp peaks. In recent years much expertise and ingenuity has been expended in constructing equipment capable of producing such curves at high resolution so that the fine structure of a melting curve occurring over 10 °C or so could be resolved into a dozen or more sharp peaks each a degree or so in width.

Wada and Tachibana, in a paper in *Nature* (**263**, 439; 1976) note that short DNA species do indeed manifest such fine structure in their differentiated melting curves and, moreover, that the melting temperature of the individual peaks does indeed bear the expected relation to the base composition of the corresponding DNA region as revealed by spectral analysis. In longer DNA species, presumably consisting of many homostable regions with closely similar stabilities, the fine structure is less apparent.

These authors then go on to consider the behaviour in more detail of a short DNA (fd phage, 6,000 base pairs) and fragments produced from it by the restriction enzyme R*Hin*HI.

First, the intact DNA (RF linear) is shown to have a differentiated melting curve consisting of more than a dozen sharp peaks distributed over a 10 °C melting range. The product formed by the action of the restriction enzyme and consisting of three fragments gives a broadly similar curve. When the three fragments were separated by polyacrylamide gel electrophoresis they individually gave curves with considerably fewer peaks. Nevertheless, when the three curves corresponding to the three fragments were summed, the result was closely similar to the curve of the mixture.

These results suggest that the peaks in the differentiated melting curves arise from homostable regions situated at specific locations in the molecule. It would appear from the relative sizes of the peaks and from the fact that splitting is observed that some peaks correspond to several homostable regions with closely similar stabilities.

A detailed comparison of the curves for the intact molecule and for the mixture of its fragments suggests that one homostable region may affect the melting temperature of others so that peaks in one of the curves may be split in the other by moving to slightly different temperatures. It is tempting to speculate that such interactions are between adjacent homostable regions, in which case it might be possible to map the regions along the molecule. Certainly the effect presents interesting statistical mechanical problems.

Wada and Tachibana end by speculating as to whether these homostable regions have some biological role. The data suggest that each operon can contain only a few and perhaps only one homostable region, and the authors suggest that the regions might play some part in recombination. They also note that wobble interactions in translation could well mean that homostable regions and active polypeptide sequences could be selected for simultaneously. □

Photosynthetic prokaryotes

from Venetia A. Saunders

The second international symposium on photosynthetic prokaryotes and first meeting of the Federation of European Microbiological Societies was held at the University of Dundee, Scotland, on August 22–28, 1976. It was organised by G. A. Codd, and W. D. P. Stewart of the Department of Biological Sciences, University of Dundee.

MUCH interest in the photosynthetic prokaryotes derives from their evolutionary and ecological significance. Moreover their metabolic versatility makes them particularly attractive research tools for studying processes involved in photosynthesis and nitrogen fixation. It is customary to recognise two groups of photosynthetic prokaryotes, the photosynthetic bacteria and the cyanobacteria (blue-green algae). The cyanobacteria typically exhibit oxygenic photosynthesis with water as the ultimate electron donor. On the other hand photosynthetic bacteria do not evolve oxygen, but utilise oxidisable substrates besides water (notably reduced sulphur compounds or organic compounds) as electron donors in photosynthesis. Now it seems that certain blue-green algae can grow at the expense of reduced sulphur compounds. Are these compounds also utilised by blue-green algae as electron donors in photosynthesis? R. Castenholz (University of Oregon) reported that sulphide could act as a photoreductant in certain blue-green algae from hot springs, but only when photosystem II was switched off. However it is not yet clear whether this process can be used to support growth. H. Utkilen (University of Oslo) demonstrated oxidation of thiosulphate by the blue-green alga *Anacystis nidulans*. He proposed that when the supply of ATP or of electrons from water is curtailed by low light intensity, *A. nidulans* utilises thiosulphate as an electron donor in a manner analogous to that of photosynthetic bacteria.

The primary photochemical reaction in bacterial photosynthesis involves transfer of an electron from bacteriochlorophyll to the primary acceptor. In the purple non-sulphur bacterium *Rhodopseudomonas sphaeroides* this occurs in a pigment-protein complex, the reaction centre. Recently intermediate short-lived states of this primary event have been identified. Interpretation of the nature of such states has been aided by picosecond

kinetic measurements. R. Cogdell (University of Glasgow) and colleagues propose that both bacteriochlorophyll (magnesium porphyrin) and bacteriopheophytin (hydrogen porphyrin) have a role in the primary photochemistry (see *Proc. natn. Acad. Sci. U.S.A.*, **72**, 2251; 1975). F. Reiss-Husson (Gif-sur-Yvette) described a procedure for incorporating reaction centres from *R. sphaeroides* into lecithin liposomes. Such an *in vitro* system should be convenient for studying the contributions of various electron transport components to the reconstitution of photosynthetic activity.

In photosynthetic bacteria the light-harvesting pigments are either accommodated with the energy conversion system on a topologically complex cytoplasmic membrane, or located in different structures of non-unit membranes called chlorobium vesicles. R. C. Fuller and C. Boyce (University of Massachusetts) reported that in the green bacteria reaction centre, bacteriochlorophyll *a* but not the bulk light-harvesting (antenna) chlorophyll is associated with the cytoplasmic membrane. The antenna chlorophyll is restricted to the chlorobium vesicles, which are considered to press up against, but remain distinct from, the cytoplasmic membrane. Light energy is transmitted from these vesicles to the reaction centre. *Chloroflexus* the only known organism with chlorobium vesicles capable of growing aerobically in the dark as well as anaerobically in the light, should accordingly prove an attractive system for investigating the development of this type of photosynthetic apparatus.

Nitrogen fixation has been more extensively studied in the cyanobacteria than in the photosynthetic bacteria. Furthermore the ability of cyanobacteria to fix nitrogen may explain symbiotic associations of certain blue-green algae in lichens and liverworts. Morphological and physiological modifications of the blue-green algae accompany these associations as reported by W. D. P. Stewart and colleagues (Dundee University). Significantly such symbiotic blue-green algae fix nitrogen more efficiently than in their free living state. The heterocyst is the main site of nitrogen fixation in certain filamentous (heterocystous) blue-green algae. Neither CO_2 fixation nor photosynthetic oxygen evolution occurs in heterocysts, such activities being restricted to vegetative cells. E. Tel-Or (University of Dundee) proposed that the defects in the photosynthetic machinery of heterocysts were probably due to the low level of ribulose diphosphate carboxylase, the cardinal enzyme in CO_2 fixation, and depletion of manganese required for

the water splitting reaction of photosynthesis.

Heterocysts develop when blue-green algae are grown in the absence of combined nitrogen. A definite heterocyst pattern emerges throughout the algal filaments. The development of such patterns may well be dependent not only on intracellular but also intercellular reactions. J. C. Meeks (Michigan State University) suggested that some diffusible product of mature heterocysts, possibly glutamine or a derivative, may inhibit nearby cells from differentiating into heterocysts. However the heterocyst pattern is established before nitrogen fixation occurs. What then triggers the developmental events? Studies on the mechanisms of induction of enzymes pertinent to the process of differentiation, as pointed out by R. Haselkorn and N. Woods (University of Chicago), may further elucidate this problem.

The promise of the genetics and molecular biology of photosynthetic prokaryotes is gradually being fulfilled. A novel gene transfer system has been discovered in *Rhodopseudomonas capsulata* by Marrs, (see *Proc. natn. Acad. Sci. U.S.A.*, **71**, 971; 1974). Furthermore this system has been used for mapping genes concerned with photopigment synthesis (B. Marrs, St. Louis University) and also for manipulating genes for nitrogen fixation (J. D. Wall and H. Gest, Indiana University) in this organism. The ability to manipulate genes in photosynthetic prokaryotes will lead to a more profound understanding of the control of such processes as photosynthesis and nitrogen fixation.

In discussing the evolution of the cyanobacterial genome, M. Herdman (University of Dundee) reported a concomitant increase in genome size with increasing morphological and biochemical complexity for these organisms. He speculated from the results that genomes of the cyanobacteria could have formed by multiplication and divergent evolution of copies of a small ancestral genome.

R. P. Ambler (Edinburgh University) reporting sequencing studies on *c*-cytochromes of eukaryotic and prokaryotic photosynthetic systems, urged caution in the interpretation of phylogenetic relationships based on structures of existing proteins. Indeed differences and similarities between species may well have been blurred by the agency of genetic exchange.

A fitting finale to the meeting was the fascinating report by R. A. Lewin (University of California, La Jolla) of prokaryotic "green algae/bacteria". These unicellular algae, found in association with ascidians, possess a unique combination of characteristics, predominantly prokaryotic, but the organ-

isms contain both chlorophyll *a* and *b* as found in eukaryotic plant cells (see *Nature*, **261**, 697; 1976). Could these organisms or relatives be ancestral forms of photosynthetic plastids of eukaryotic cells? Whether this type of "prokaryote" is a missing link or not remains to be seen, but it is likely to be a link between this and the third symposium. ☐

Anti-juvenile hormone and pest control

from a Correspondent

THE preparation of active extracts of juvenile hormone from the American cecropia silkmoth 20 years ago suggested the possibility of the eventual use of synthetic hormone as an insecticide. The original idea was that this could lead to the appearance of non-viable intermediates between larvae and adults. The discovery, 10 years later, of the chemical nature of juvenile hormones (of which three varieties have so far been described) stimulated an immense amount of synthetic chemistry in this field. A vast range of products has been produced, some of them not at all closely related chemically to the natural hormone; some, far more active than the natural substance, have been put on the market as insecticides.

The obvious theoretical objection to the juvenile hormone as an insecticide —that it will tend to produce giant larvae which may be more destructive than the ordinary sort—has proved justified in some cases. On the other hand, it has been proved empirically that some of these products are effective for rather special purposes, and some of them have a sterilising action on the adult female or disrupt the development in the resulting eggs. By and large, however, the results have been somewhat disappointing.

It was suggested early that an anti-juvenile hormone, which could cause precocious metamorphosis, might be a more successful kind of insecticide. In a recent paper William S. Bowers and his colleagues at Geneva, New York (*Science*, **193**, 542–547; 1976) report a search for such products in a long series of plants, by extracting the less polar organochemical components in a mixture of ether and acetone. From the common bedding out plant *Ageratum houstonianum* they obtained a product which had very active anti-juvenile hormone properties when applied to immature Hemiptera. It in-

duced precocious metamorphosis at the ensuing moult when applied to first, second or third-stage larvae of the milkweed bug *Oncopeltus;* and similar results were obtained in the cotton stainer *Dysdercus.*

The active principle was purified, analysed and synthesised. It contained two simple chromenes, 7-methoxy-2, 2-dimethyl chromene and 6,7-dimethoxy-2, 2-dimethyl chromene (Fig. 1), which were already known from plant extracts. On account of their ability to cause precocious metamorphosis these substances have been named precocene 1 and 2. Precocene 2 is about ten times more active than precocene 1.

Precocene 1   Precocene 2
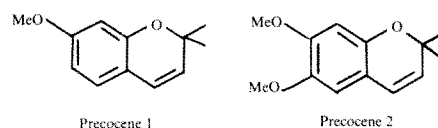
Eggs of Hemiptera fumigated with the precocenes gave rise to larvae which developed normally for two instars and then moulted to precocious adults. The precocious adults, however produced, had well formed ovaries but they never matured to form ripe eggs —a process for which juvenile hormone is necessary. Application to adult females caused regression of the ovaries; and late treatment had an ovicidal action, the nature of which is unknown.

A hundred years ago

The Colorado potato beetle *Leptinotarsa*, on the approach of winter, ceases to secrete juvenile hormone. It enters reproductive diapause, egg development is suppressed and it burrows into the soil. When the adult *Leptinotarsa* was treated with precocene it was forced, apparently permanently, into diapause. The mode of action of the precocenes is not yet known; there are many steps in the synthesis, physiological action or breakdown of the juvenile hormone, any of which could be responsible for the disruptive effects observed.

It remains to be seen whether materials of this kind will prove effective for practical control. The fact that all their effects will diminish insect damage, and that all the precocious females are sterile, are favourable factors. The substances obtained so far have induced precocious metamorphosis only in Hemiptera, and not in Holometabola. But they have sterilised adult females among Diptera and Coleoptera. The discovery of anti-juvenile hormones specific for selected pest species is an atractive possibility. ☐

Specific toxin receptors in plant disease

from I. M. Smith

THE fungi which parasitise plants are commonly host-specific; one pathogenic species will often only infect a limited range of closely related plant species or even only some genotypes within a species. The molecular basis of this specificity is one of the outstanding problems in physiological plant pathology. Some progress has been made in the type of disease in which successful infection depends on the release by the fungus of phytotoxic metabolites which may be host specific. One notable example of this is the Victoria blight of oats (caused by *Helminthosporium victoriae*), in which the toxin victorin has high specific activity only on certain oat genotypes, which are at the same time the only ones susceptible to infection. Similarly, *H. sacchari* produces the toxin helminthosporoside which is only toxic to susceptible lines of sugarcane. In this case, Strobel has reported a membrane protein which binds the toxin, but only in susceptible lines. Resistant lines have a slightly different protein which does not bind. The immediate consequences of toxin binding are not clear, but this work has been till now the only reported case of a specific receptor for a phyto-

toxin involved in pathogenesis. Now, Steele *et al.* (*Proc. natn. Acad. Sci. U.S.A.*, **73**, 2245; 1976) report the specific binding of another phytotoxin (tentoxin from *Alternaria tenuis*) to a receptor protein in susceptible hosts.

A. tenuis is a widespread weak parasite which causes near-saprophytic infections of grain, fruit and senescent tissues generally. It does, however, cause a very distinctive seedling blight of some species, in which the cotyledons become severely chlorotic and the fungus invades the damaged seedling. This chlorosis can be exactly reproduced by treatment with tentoxin, a cyclic tetrapeptide, and it seems that the fungus is dependent on the action of the toxin for successful infection.

Steele *et al.* have now shown that tentoxin shows specificity in three respects. When added to isolated chloroplasts of a toxin-sensitive species such as lettuce, it inhibits coupled electron transport. When added to purified coupling factor 1 (CF₁, a chloroplast ATPase directly involved in photophosphorylation), tentoxin inhibits the ATPase activity. If synthetic tritiated tentoxin is added to CF_1 in a continuous ultrafiltration system, the protein can be shown to have approximately one binding site per molecule for tentoxin, with an affinity constant of $2 \times 10^8 \, M^{-1}$. Comparable tests on an insensitive species (radish) showed only slight inhibition of photophosphorylation, no inhibition of ATPase activity, and an affinity constant less than $10^4 \, M^{-1}$. Other sensitive and insensitive species fitted the same pattern with respect to inhibition of photophosphorylation. It would seem, therefore, that the CF_1 of sensitive species acts as a tentoxin receptor and that toxin-binding leads to inhibition of photophosphorylation. The CF_1 of sensitive species binds poorly and is not inhibited. These results are of great significance, first because of the specificity of the inhibition and binding, and second because the binding may in itself explain the mechanism of toxicity of tentoxin.

There are, however, certain aspects that remain to be analysed. The relation of toxin specificity to host specificity is complicated by the wide host range of the fungus and the different diseases that are caused. Toxin specificity may be presumed to relate only to the specific ability to cause a certain type of disease, that is seedling chlorosis. Certain plants, such as corn, show sensitivity of isolated chloroplasts but not of intact leaves, so that other factors may operate to cause insensitivity. In addition, the symptoms of chlorosis are most strikingly seen in seedlings with developing chloroplasts. The details of the causal relationship between inhibition of photophosphorylation and

this chlorosis are not yet established. Nevertheless, tentoxin could clearly become the first disease toxin for which the molecular basis of both specificity and toxicity can be explained. □

Coming to terms with the noisy retina

from Jonathan Ashmore

An international symposium on Photoreception was held at the Royal Society on September 2 and 3, 1976. Sponsored by the Rank Prize Funds, the symposium was organised by P. Fatt and H. B. Barlow.

RETINAL physiology has come to possess a decided flavour of engineering. Touching on problems close to the heart of any communications engineer, the mechanisms used by cells of the retina to transduce light signals in the presence of intrinsic noise were a recurrent theme at this meeting.

There now exist a number of complementary studies on the first stages of visual excitation, from the absorption of a photon by the photoreceptor pigment molecules to its transduction manifested as a change in membrane potential of the cell. There is still no evidence directly in conflict with the 'Calcium Hypothesis' advanced some years ago by W. Hagins and coworkers, (*A. Rev. Biophys. Bioeng.*, **1**, 131; 1972), which postulates that in rods calcium is released by light from the intradiskal space to diffuse intracellularly and close ionic channels on the cell membrane. Equally, there is no direct evidence supporting the idea, but Hagins and S. Yoshikami (National Institutes of Health, Bethesda), however, think that calcium is still the best bet for the intracellular transmitter. At this meeting they reported experiments intended to change the intracellular ionic environment in rat rods by fusing vesicles containing chelating agents with the cell wall. Model calculations suggest that calcium is the only ion with a sufficiently high affinity in these experiments to account for the observed changes in the rods' electrical photoresponse. Direct intracellular injection experiments in the toad rod system were reported by L. H. Pinto (Purdue University), J. E. Brown and J. A. Coles (Vanderbilt University), but again produced no evidence contradicting Hagins' story.

More than half the symposium was given over to the interactions between receptors and other cells of the retina. Since the initial indirect evidence 18 months ago by G. L. Fain that toad rod receptors are coupled (*Science*, **180**, 1178; 1975), the demonstration of inter-receptor coupling in various species has become almost a minor cottage industry. W. G. Owen and D. R. Copenhagen (University of California, San Francisco) discussed double impalement experiments with microelectrodes, showing conclusively that the coupling between rods in the snapping turtle is electrical, probably by way of the telodendria extending from the pedicles, (see *J. Physiol.*, **259**, 251; 1976). The effect of this coupling is to attenuate the signal from a singly illuminated rod to 10–20 times that produced when all the receptors are illuminated. One function for the coupling, at least in the peripheral retina, would be to improve the signal-to-noise ratio for dim diffuse lights, which would be traded off against the loss in visual acuity caused by coarsening the receptor mosaic.

Intracellularly recorded noise in turtle cones, probably produced by the opening and closing of ionic channels, has been used by T. D. Lamb and E. J. Simon (University of Cambridge) to show that in this animal the cones are electrically coupled. Quantitative measurements of the noise lead them to conclude that in these cells one photon can close one ionic channel, whereas in the rod system, as E. A. Schwartz (Harvard Medical School) reported using the same power spectrum methods, a figure closer to one photon controlling 300 channels may be more in order.

A nice bridge between electrophysiology and behavioural studies was made by G. L. Fain (Jules Stein Institute, UCLA). Using data from the behaviourally determined increment threshold curve for turtle and comparing this with the intracellularly determined light sensitivity, he estimates that a 5–10 μV signal is required in red cones for a behavioural response. This figure is below the cone noise level, although many parallel pathways exist in the retina for output. The properties of some of these retinal pathways have been mapped by D. A. Baylor and R. Fettiplace (Stanford University) by injecting current from a micropipette into cones and recording the ganglion cell response. They suggest that the pathways studied in the cone system may afford some signal-to-noise improvement by functioning as a bandpass filter; in the rod system at threshold where the problem becomes more acute a similar characteristic could be present.

The ionic mechanisms involved in

synaptic transmission have proved difficult to disentangle since usually more than one cell type meets the receptor at synaptic contact. At the synapse between receptor and depolarising bipolar cells, J-I. Toyoda (Kawasaki University) describing experiments to pass current trans-retinally produced evidence that at least two postsynaptic ionic channels are transmitter modulated. The hyperpolarising bipolar cells still conform to the original suggestion that light produces a permeability decrease to a single ion species.

A method to decouple the horizontal cells from their putative role in the feedback loop at the receptor triad synapse would be very helpful here, but failing this an interesting technique was reported by Yù. A. Trifonov and A. L. Byzov (Moscow University) to study the electrical properties of the horizontal cell membrane. Because of extensive electrical coupling between these cells, microelectrode studies of the cell membrane are not feasible; the method used here is a modification of Kramer's *'Winkel rinne'*, and results suggest cell membrane possesses anomalous rectification properties. Whether this will yield further information on the receptor synapse remains to be seen.

We may now be in a position to localise the cellular origins of the 'noise' treated in the psychophysical literature as the ultimate determinant of visual discrimination. Stressing this point, however, H. B. Barlow (University of Cambridge) tentatively suggested from experiments using random dot patterns that the factors which limit the quantum efficiency of seeing determined from performance tests may not arise in the retina at all but from "nasty central processes going on". Central limitations to quantum efficiency can only revise estimates of retinal efficiency upwards. ☐

Pesticides and the environment

by Mary Lindley

A symposium on the ecological effects of pesticides, sponsored by the Institute of Biology and the Linnean Society was held in London on September 23–24. The proceedings will be published by Academic Press.

PESTICIDES are not likely to be abandoned to the accompaniment of cries of ecological doom, although they

clearly can be misused and have caused death and destruction in some cases. DDT and its successors have wrought so much good in controlling disease and increasing crop productivity that they must continue to be used, albeit in moderation and with constant efforts to find alternatives. This was the message of the symposium, during which very few speakers discussed directly the ecological effects of pesticides.

According to H. C. Gough (Ministry of Agriculture, Fisheries and Food, London), the benefits of pesticides used correctly in agriculture far outweigh their hazards. Their use has greatly reduced losses of many crops from pests, diseases and weeds. Without pesticides, he said, production would decline and costs would increase. Unnecessary use occurs, however, and there is a continuing need to ensure that hazards are minimised in the application of pesticides.

A particularly hopeful development in crop protection from the ecological point of view is selectivity. I. J. Graham-Bryce (Rothamsted Experimental Station, Harpenden) cited in particular the pyrethroids, which are potent insecticides, through their actions as nerve poisons, but not harmful to mammals. Another target for selective pesticides is the endocrine system of insects; juvenile hormone analogues and anti-juvenile hormone compounds hold promise.

The control of pests and diseases in forests presents a greater dilemma than when agricultural crops are concerned. Pesticides are in any case more likely to upset biological equilibria in perennial than in annual plant environments and the multipurpose nature of many forests adds to the complications, as M. J. Way (Imperial College Field Station, Silwood Park, Ascot) explained. He pointed out that the balance between species in forests may indeed be maintained by pests: for example, without spruce budworm some forests would be monocultures of spruce. But on the whole, pesticides are used less in forestry than in other major production systems.

The disastrous ecological consequences of the use of herbicides to defoliate large tracts of forest in Vietnam during the recent war were described by Dr A. H. Westing (Stockholm International Peace Research Institute). Severe destruction of habitats was caused in more than 50,000 hectares of upland forest and in 150,000 hectares of mangrove forest during the American programme of spraying with 2,4-dichlorophenoxyacetic acid and 2,4.5-trichlorophenoxyacetic acid. According to Westing, the damage will be long lasting and has effectively deprived Vietnam of a proportion

of its forests.

The difficulties of monitoring the ecological effects of pesticides, particularly in developing countries, were made very clear by W. E. Kershaw and M. Pugh Thomas (University of Salford). They have been studying the effects of insecticides used to control the blackfly vector *Simulium damnosum* in the World Health Organization Onchocerciasis (river blindness) Control Programme. They have found that investigations of the biological side effects of pesticides on fast flowing rivers in Africa are expensive, tedious and hazardous. So far they have no evidence that fish have been affected, while the data obtained for invertebrates are very hard to interpret. Pugh Thomas believes that to allow for seasonal effects it will be necessary to compare data for successive years before a clear picture is likely to emerge.

The problems involved in predicting how pesticides will affect the environment were underlined by F. Moriarty (Institute of Terrestrial Ecology, Monks Wood Experimental Station, Huntingdon). Direct field trials are difficult, and the various indirect approaches used have disadvantages. But he said two tentative assumptions are possible: organisms have little effect on the amounts of pesticide in the physical environment, and if individual organisms are affected adversely, then ecological effects are imminent or already happening.

The prospects for wildlife, according to N. W. Moore (Nature Conservancy Council, London), depend more on the future of their habitats than on the use of pesticides. The decline in the wildlife population in Britain during the past 150 years has been due more to the destruction of habitats for land development than to the effects of pesticides. He pointed out, however, that treatment with pesticides can be locally damaging, and it will be increasingly important to ensure that pesticides do not spread away from the crops which they are protecting, for example by drift or in effluent. He said that the future of wildlife in the farmed countryside will depend on the conservation of areas of unfarmed habitat, where pesticides cannot reach.

Discussing future prospects for man, K. Mellanby (Institute of Terrestrial Ecology, Monks Wood Experimental Station) said that although in an ideal world pesticides would be unnecessary, they will continue to be used, and with proper precautions and controls, they can do their job while man's environment continues to be improved. He said that the dangers inherent in the use of pesticides are generally accepted and there has often been public overreaction to them. ☐

articles

Astronomically-oriented markings on Stonehenge

Richard F. Brinckerhoff

Science Department, Phillips Exeter Academy, Exeter, New Hampshire 03833

In view of the years of careful study that have been devoted to Stonehenge it may seem unlikely that any more information could be extracted from the monument, or at least from that part of it above ground. During the past two years, however, further markings have come to light that are, at least, interesting, and possibly important. These markings are a series of at least 11 pits on the upper surfaces of the three contiguous lintels (130, 101 and 102) that span the well-known line of sight from the centre of the sarsen circle north-eastward towards the heel stone. For an observer diametrically across the circle, 9 of these pits identify directions of the rising moon at significant points in its 18.6-yr cycle.

THE lintel-top surfaces of Stonehenge maintain a true level with the horizon to within a few inches in spite of the fact that within the sarsen circle the ground level drops down towards the north-east by ~ 1.5 feet. If the sole purpose of the lintels was architectural, it is hard to see why such a small correction-to-level would have justified the effort; its architectural effect would be imperceptible. But as a walkway for observing horizon events such a correction would make sense. Moreover, the diminutive, upright stone 11 in the south-east quadrant could then be explained as a midway support to a stair which, forking to left and right at its top, gave access to the lintel-top walkway.

This hypothesis suggests the possible existence of astronomically oriented markings on the lintel surfaces, but nowhere in the Photographic Library of the Department of the Environment was there a close-up picture of any of the nine lintels now in place. Direct inspection was therefore necessary.

In July of 1974 with the permission of the Department of the Environment and the help of Mr Woodhouse, the head custodian, a ladder was provided and all nine lintels

inspected. Only lintels 130, 101, and 102 (the three contiguous ones in the north-east quadrant of the sarsen circle) showed promising markings[1]. In 1975, again with permission and help, a series of measurements, photographs, and plasticene moulds were carefully made. Figure 1 shows the surface of the three lintels made by piecing together a sequence of photographs made approximately every 2 feet along the lintel tops from a height of 5.5 feet. They were taken with the Sun low in the west, at about 1800 on June 13, 1975, shadowing the rock surface. The measuring stick is 2 feet long.

Pits

The significant markings are a series of pits, some of which in the photograph contain vertical wands used for locating them precisely from the ground. The pits are identified by numbers in Fig. 2, which is made from a tracing of a photograph, originally 2.5 feet long.

Pit 1, the largest, is almost circular both in plan and in section, ~ 10 inch across and 3 inch deep. (The numbers here given are only approximate because of the irregular surface of the lintels, and the fact that not all the pits have clearly defined edges.) Pits 2, 3 and 4 are much smaller, 1.5–2 inch in diameter, quite round but irregular in depth. Pit 5 is round, 3 inch in diameter and about as deep. Near it is a pit, 6, that may be more significant though it is of irregular shape, an uneven 3×1.5 inch in plan and up to ~2 inch deep. Pits 7 and 8 command attention because of their location almost directly over the sight line from the sarsen circle centre to the heel stone. They are both approximately 4 inch in diameter and depth with regularly rounded bottoms. Finally, pits 9, 10 and 11 are shallow, smaller, and not quite so regular in shape. Pits 1 to 8 (excluding 6) are so regular in shape as to be plausibly man-made; the others together with several smaller markings command less confidence, but all of them stand out sharply in appearance from the uneven weather-worn surface of the lintels. Indeed, pits 1, 7 and 8 can be detected in some low level air views already in print.

Fig. 1 Composite photograph of lintel tops from a height of ~ 5.5 feet showing pits. The measuring stick is 2 feet long.
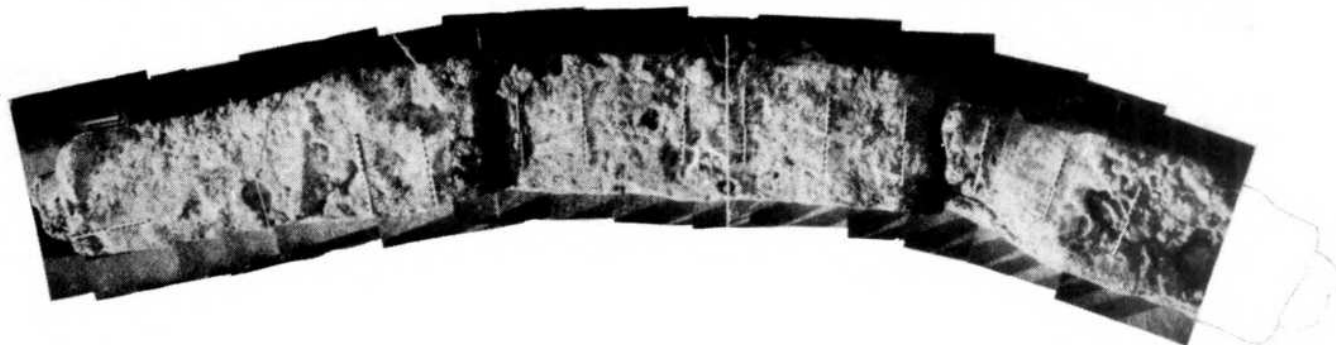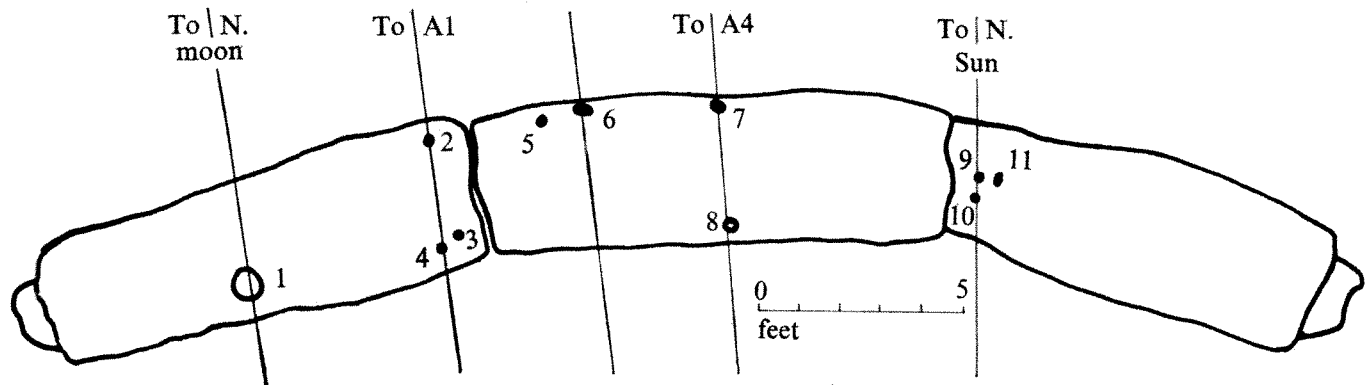
Fig. 2 Identification of pits in Fig. 1 and the alignments through them. The directions shown are approximate.

Alignments

To pursue the possibility that the pits are man-made it is necessary to locate them accurately on the overall plan of the monument. It was found by projecting Hawkins's 1965 photogrammetric air survey map[2] on to an enlarged drawing of the Thoms' more recent survey[3] that the two agree precisely in all essentials. We used the air survey, except for the replacement of the three lintels by a reduced rendering of Fig. 2.

Nine of the eleven pits identify significant directions for an observer posted on the now non-existent lintel top diametrically across the sarsen circle at P (Fig. 3). P falls on the centre of sarsen upright 15.

When the lintel-top pits are plotted on the air survey map it is found that a line through pit 1 to the northernmost point of moonrise and a line through pits 9 and 10 to the northernmost (midsummer) sunrise intersect on the south-western side of the sarsen circle at P. P is defined by the intersection of these two sight lines.

Moreover, as viewed from P, pits 2 and 4 point very nearly but not precisely to the northernmost of the four A holes (A1). Similarly, a line from P through pit 5 (not drawn in Fig. 3) points very nearly to A2. A line from P through the Aubrey hole centre shown as a dot (but not through the sarsen circle centre shown as ✕) passes through pit 6 and very close to the midpoint of the A holes between A2 and A3. And the conspicuous pits 7 and 8 over the central archway point very nearly to A4. All of these lines miss the A holes by the same amount and in the same direction (eastward), a matter discussed further below. Only pits 3 and 11 elude an obvious pattern. So simple a pattern

fitted so precisely by such a large majority of the pits is additional evidence that the pits (or at least nine of them) are man-made.

From point P none of the pits aligns with the rising point of Venus or of any of the 20 brightest stars except, implausibly, Pollux after corrections are made to the year 2000 BC for the precession of the pole. Nor do any of the pits seen from the sarsen circle centre or the Aubrey hole centre align with any significant sunrise or moonrise points or with Venus or the brightest stars.

A shift in the observation point

There is a strong intuitive appeal in the idea that the viewing point should not be at P but at a point ~5 feet westward so that the observer would be on what the Thoms[3] have identified as the axis of symmetry of the monument. There he would have a clear view under the lintel of the great trilithon whose bottom surface would be 5.5–6 feet above his footing atop the sarsen circle. The angular spread of the alignments (pit 1 to pits 9, 10) is 10°, too great to be seen through the trilithon from one spot, but as the observer walked north-westward from P along the lintel tops he would see the pits revealed progressively through the arch as he moved. Pit 1 would come into view (azimuth 42.3°) in approximately the direction of the northernmost moonrise (azimuth 40.5°), and pits 7 and 8 would be seen almost (but not quite) in line over the sarsen circle centre, with the centre of the heel stone at 51.7°; the azimuth of the midsummer sunrise is 50.7°. These are unsupportably large discrepancies. Moreover, viewed through the trilithons in this way pits 2–6 and 9–11 have no obvious significance.

Fig. 3 The alignments through the lintel-top pits to outlying A holes and horizon points.
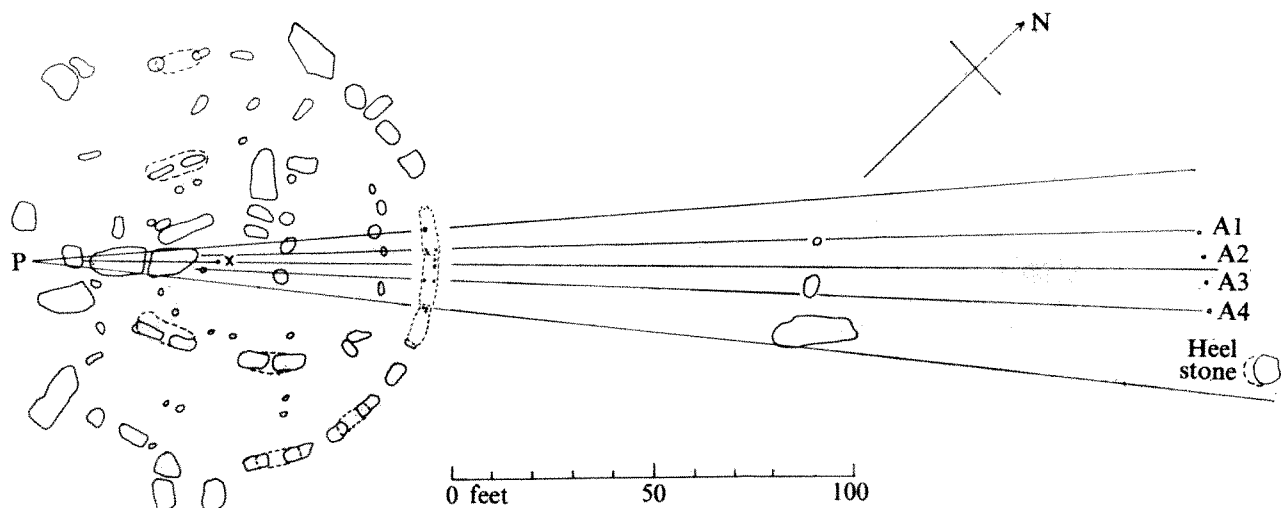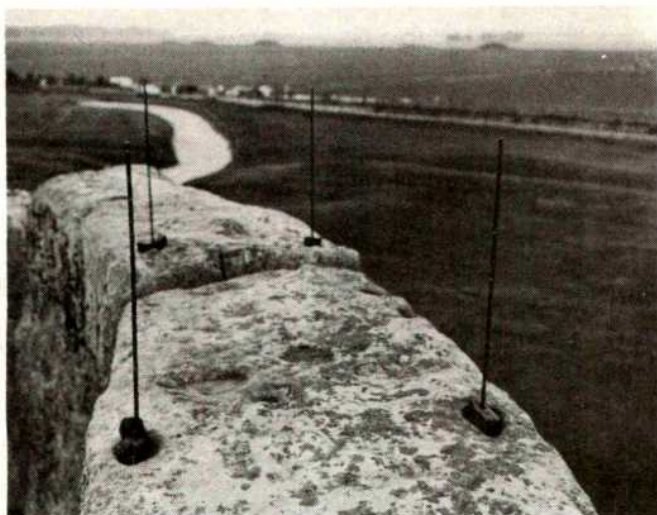
Fig. 4 Close-up of some of the pits showing how they may have been used to support wands as sights.

Also, when vertical wands are erected in pits 2, 4 and 7, 8 (Fig. 4) they appear precisely aligned only as seen from within ~ 1 foot from P (Fig. 5) and not through the trilithon arch, 5 feet from P. Finally, since the axis of the monument is aimed very nearly at midsummer sunrise, it can scarcely be aimed simultaneously at, or be symmetrical with, moonrise points all to the N. On the whole, the alignments radiating from P are impressively accurate.

Orientation

With the help of plumb lines pit 1 was located at ground level on the monument plan to within 4″, and its azimuth was found to be 40.5°. Using the procedure given in ref. 4 the first gleam of the northernmost moonrise ($\delta = +29.1°$) on the horizon of Stonehenge is found to have an azimuth of 40.7°.

Similarly, the ground level location of pits 9 and 10 identify an azimuth of 51.0° from P, whereas the first gleam of midsummer sunrise ($\delta = +23.9°$) seen over the heel stone from the sarsen circle centre has an azimuth of 50.1°.

With pits 1, 9 and 10 thus located on the plan, a tracing of the lintel-top photograph (Fig. 2) was reduced in size until pits 1, 9 and 10 on the photograph registered accurately with the plan. The lintel-top tracing was then drawn, thus locating accurately all the intervening pits. So located, the pits identify the lines toward the A holes (Fig. 3) (See Table 1 for azimuths).

The accuracy with which the sunrise and moonrise lines are located on the map is limited by uncertainties of an inch or two in the restoration of the three lintels in 1919–20 and the fact that location of the pits on the map are accurate to only ~ ±4 inch and that angles measured in Fig. 3 are no better than ±0.25°. Together these errors result in an uncertainty in the location of P of ~ 1.0 foot.

Accuracy

In an effort to judge the accuracy with which the lintel-top pits align with the A holes, a camera was set up approximately at point P. In Fig. 5 beneath the right-hand lintel top wands in pits 7 and 8 is an arrow pointing to a post erected in the centre of hole A4, almost exactly three sarsen circle diameters from the camera. Although pit A1 is obscured it can be located in effect by simple geometry 20.4 feet left of A4, which agrees with the distance actually measured on the ground. Measurements on the photograph show that to make the pits align with the A

holes, the camera would have to move 1.0 foot eastward to be centred on P and then an additional 1.4 feet eastward beyond P. This is appreciably more than the ±1.0-foot uncertainty, inherent in the original map location of P. Thus, in spite of their identity of spacing, the set of sarsen circle alignments appears to be shifted slightly to eastward of the set of A-hole alignments. It looks as if the builders had come up with two slightly different solutions to the same observational problem. For an explanation of this mismatch we must first examine the consequences of the fact that the A holes and the sarsen circle lintel pits were in use at two widely different times.

Stonehenge I

The A holes 3 feet in diameter, were only 4 feet deep, much too shallow to support the 28-foot posts required for a man standing on the lintels to use them for sighting. Moreover, post hole A1 was not only filled in but was entirely buried by the bank created by the building of Stonehenge II. Together these facts point to the use of the A-hole posts as sights by ground-level observers before the construction of Stonehenge II.

At that early time, when there was no sarsen circle, it seems likely that the observers were stationed at the centre of the Aubrey hole circle rather than at P. The reason for this is simply that from the centre of the Aubrey hole circle the four A-hole posts (centre line azimuth 45.5°) would have been centred between the line to the heel stone (50.7°) marking the northernmost sunrise, and the line to the northernmost moonrise (40.5°). Posts on the A holes would then mark the 18.6-yr swing of the moon during intermediate years. When the Moon rose through the first opening (A4–A3) of the four posts, it was just one quarter of the way in time from rising on the line of midsummer sunrise and rising at its northerly extreme; when it rose through the middle opening it was just one third of the way to its maximum, and when it rose behind A1 it was exactly halfway.

Although the precise date of the maximum would be impossible to identify accurately, just as for the Sun, since, close to the maximum, the change in azimuth with time is too small to observe, the date could nonetheless be inferred quite accurately in retrospect as lying halfway in time between the northbound and the southbound traverses of any one of the three A-hole openings. After one or two such pairs of traverses (18.6 yr apart) the observers could confidently predict the proper date for ritual celebrations of the maximum.

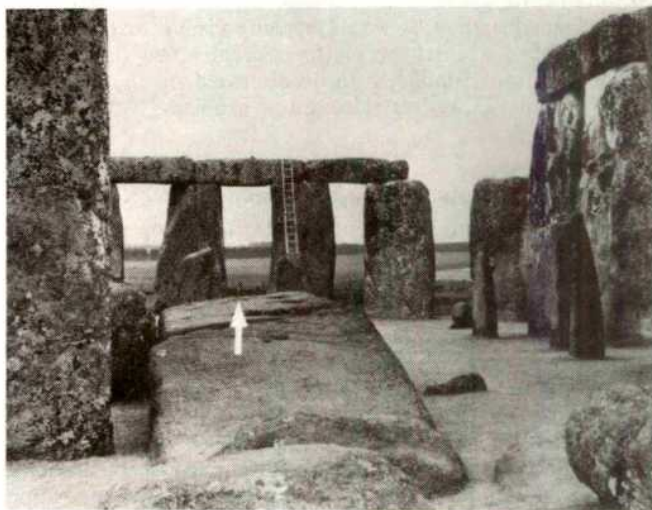**Fig. 5** Ground-level view of the wands in Fig. 4 from a point 1 foot west of the presumed position of P.

Table 1 Orientation data of the pits on Stonehenge lintels

Alignment from P through pit:	Towards	Azimuth E of N±0.2°	Vertical error, D°, for		
			First gleam	Centred on horizon	Tangent on skyline
1	Northernmost moon $\delta = +29.1°, A = 40.5°$	40.5	−0.09	−0.39	−1.67
2, 4	A1	43.5			
3		43.6		> 1.5	
5	A2	44.5		for both Sun and Moon	
6	Mid-A and Aubrey centre	45.5			
7, 8	A4	47.5			
9. 10	Midsummer sunrise $\delta = +23.9°, A = 50.7°$	50.7	+0.63	+0.40	+0.10
11	Midsummer sunrise	51.2	+0.92	+0.68	+0.45
Through centre of great trilithon to pit:		**Azimuth E of N ± 0.6°***			
1	Northernmost moon $\delta = +29.1°, A = 40.5°$	42.3	+0.75	+0.52	+0.29
2, 4		45.8		> 1.5°	
3		47.5		for both Sun and Moon	
5		47.8			
6	Midsummer sunrise	48.5	−0.53		
7	Midsummer sunrise	51.7	+1.18	+0.94	+0.82
8	Midsummer sunrise $\delta = +23.9°, A = 50.7°$	51.8	+1.23	+0.99	+0.87
9, 10		56.0			
11		56.5		> 1.5°	

*Error is higher than from P because of uncertainty concerning midpoint of trilithon arch.

The end came when the Beaker invaders, bringing with them a new culture, realigned the axis by the construction of Stonehenge II, evidently focusing on the midsummer sunrise and burying the A holes. The Moon seems to have been replaced by the Sun in the culture of the newcomers.

Stonehenge III

Centuries after the A holes had been obliterated the sarsen circle was erected, its axis also oriented to the midsummer sunrise, as was Stonehenge II before it. Then or later, wands supported upright in the new lintel-top pits must have once more identified the old alignments for observers at P, now 16 feet up off the ground to avoid being blocked by the new sarsen circle and the old bluestone circle uprights.

In their paper[5] on Stonehenge as a possible lunar observatory the Thoms discuss the improvement in visibility of foresights on distant hills made possible by raising the level of the observer above the ground level. Is this consideration one of the determinants of the height of the lintel-top walkway? No one has ever explained convincingly why it had to be 16 feet high. Do our pits identify distant foresights visible only to a lintel-top observer? The Thoms mention C. A. Newham's[6] observation that if Stonehenge were raised only a few feet, distant hills would become visible to the northeast. Certainly a lintel-top platform creates new possibilities for observations similar to those discussed by the Thoms[3,5] and Newham[6] as well as raising new problems.

Difficulties

One problem that has a plausible solution is the mismatch between the A-hole alignments and the lintel-top pits alluded to earlier and illustrated (and exaggerated) in Fig. 5. Their alignments shifted slightly eastward, the lintel-top pits appear to attest to the slow eastward drift of the extreme declinations of the Sun and Moon caused by the slowly changing obliquity of the ecliptic. For example, in 3000 BC the maximum declination of the moon was 29.18°, and in 2000 BC it was 29.08° (ref. 7). As a consequence the northernmost rising point of the Moon shifted 0.21° eastward. But a 0.21° shift eastward corresponds to the observer moving 0.5 feet eastward from the presently plotted position of P. This, together with the 1.0-foot uncertainty in P's

location amounts to 1.5 feet, a shift in P sufficient to bring the two sets of alignments into precise coincidence, assuming the maximum possible error in the original location of P. (A 1.4-foot shift in the map position was required.) In principle, if the position of P were known exactly and our interpretation is correct, the angular difference between the two sets of alignments would lead to a knowledge of the time difference between Stonehenge I and the creation of the lintel-top pits on Stonehenge III.

A much more serious problem remains. When Stonehenge III had been completed stone 55 of the great central trilithon, now fallen, would have completely blocked the view of the lintel-top observer at P. Perhaps the trilithon was erected only after the sarsen circle and the pits had been in use for some time, but this seems unlikely since it would be almost impossible to erect the trilithon after the sarsen circle had been completed. Nor do our pits seem to identify any useful alignments to an observer looking through the trilithon arch.

Speculations

Perhaps in the troubled times towards the end of the use of Stonehenge as an observatory, the trilithon was pushed over —to make way for the re-establishment of the lunar alignments—and has lain broken in its present position ever since. If this event is archaeologically dubious, it at least has the merit of accounting for most of the pits.

More plausibly, perhaps, a 4-foot high wooden platform may have been erected at P to permit viewing over the top of the trilithon.

In conclusion, although we lack an exact explanation of how sightings could have been made or why the alignments were off-axis, these results nevertheless have significance in our understanding of the astronomical aspects of Stonehenge. The original theory that the site was an observatory accurate to the width of the archways is now augmented with alignments accurate to within a few inches.

1 Hawkins, G. S., in *The Encyclopaedia Brittanica Yearbook* (1976).
2 Hawkins, G. S., *Beyond Stonehenge*, 59 (Harper and Row, New York, 1973).

3 Thom, A., Thom, A. S., and Thom, A. S., *J. Hist. Astr.*, 5, 75 (1974).
4 Hawkins, G. S., *Beyond Stonehenge*, 291pp (Harper and Row, New York, 1973).
5 Thom, A., Thom, A. S., and Thom, A. S., *J. Hist. Astr.*, 6, 19 (1975).
6 Newham, C. A., *The Astronomical Significance of Stonehenge* (Leeds, 1972).
7 Hawkins, G. S., *Beyond Stonehenge*, 295 (Harper and Row, New York, 1973).

$^{207}Pb/^{206}Pb$ whole-rock age of gneisses from the Kangerdlugssuaq area, eastern Greenland

W. P. Leeman & E. J. Dasch

Department of Geology, Oregon State University, Corvallis, Oregon 97331

M. A. Kays

Department of Geology, University of Oregon, Eugene, Oregon 97403

We present here a brief description of geological relations and results of lead isotopic analyses of eight whole-rock samples of Precambrian gneiss from an area near the Skaergard Intrusion in the Kangerdlugssuaq region of eastern Greenland.

EXCEPT for Wager's[1] regional reconnaissance studies, there have been few published details of the Precambrian geology of Eastern Greenland. Berthelson and Noe-Nygaard[2] have suggested that basement rocks of the Kangerdlugssuaq region belong to the Kenoran tectonic division (that is, rocks of pre-Ketilidian, or more than about 1,800 Myr old) of the Canadian–Greenlandic Shield. A limited number of K–Ar ages[3-4], ranging from 2,700 to 1,600 Myr BP, have been reported for Kangerdlugssuaq area gneisses, but interpretations of these data are suspect because of widespread Tertiary igneous activity in the area[5-7]. Our field studies have indicated a complex history for the gneisses, which is summarised in Table 1.

Interlayered mafic and felsic gneisses represent the earliest recognisable unit in the area (event 1, Table 1). These interlayered gneisses are interpreted as dominantly metaigneous rocks, but a small amount of intercalated metasedimentary material is locally present. Strong banding is developed in the gneisses primarily among quartzofeldspathic units. Distinctive mafic bands a few centimetres to several metres thick are also present. This gneiss suite has characteristics similar to those described for gneisses in western Greenland and equivalent gneisses elsewhere in the Archean Craton of the north Atlantic region[8-10]. In the Kangerdlugssuaq area mafic gneisses vary from nearly pure hornblendite to plagioclase-rich amphibolite. Locally, this lithology is abundant, forming mappable units of mafic gneiss up to 200 m thick. More commonly, the mafic gneisses have been pulled apart, rotated, and fragmented within a matrix of quartzofeldspathic gneiss, giving the impression that large mafic gneiss xenoliths have been rafted and deformed by later emplaced felsic gneisses. This impression may, however, be an oversimplification in as much as the mafic gneiss blocks are in places definitely interlayered with quartzofeldspathic to biotitic gneisses. Thus, it seems probable that some blocks of felsic gneiss also occur as xenolithic rafts.

Metasedimentary rocks which occur among the early gneisses are dominantly felsic gneisses with variable amounts of plagioclase, alkali feldspar, and quartz and accessory biotite and hornblende. With increasing quartz content these gneisses grade to sugary textured quartzites; with increasing biotite they become schistose and some contain garnet porphyroblasts in a fine-grained quartz and feldspar matrix. Hornblende-rich and/or biotite-rich layers are also present in the felsic metasedimentary gneisses. It

is possible that these metasediments are remnants of a once extensive supracrustal sequence that was incorporated into the metaigneous suite either as conformable sedimentary beds in a volcanic sequence or as infolded units. Alternatively, they may reflect conformable tectonic interleaving, as inferred for the Malene supracrustals within the Amitsoq gneisses in Godthabsfjord[8].

Textural and structural relationships within the later sequence of metaigneous rocks (event 2, Table 1) suggest that there were several intrusive pulses. The early intrusive rocks are commonly granite or trondhjemitic. Because of their similarity in composition and texture to felsic rocks of the earlier interlayered gneisses, contacts between the two units are discernable only in a few favourable localities. Discordant contacts are most clearly displayed where rafted blocks of agmatised interlayered mafic–felsic gneiss are enclosed by intrusive granitic or trondhjemitic gneisses. Locally, the metaigneous gneisses display pronounced coarsening, largely conformable with foliation, and associated with the occurrence of pegmatitic patches and lenses; in extreme cases these zones become essentially migmatitic. It is not clear whether the coarsening effect and migmitisation are related to a separate intrusive pulse. In some areas granitic rocks occur with little internal fabric; these bodies are apparently the latest recognisable intrusions in the area.

The grade of metamorphism throughout the entire gneiss sequence is amphibolite facies, except in areas where migmatisation is most widespread and/or intrusive rocks are more abundant. There, pyroxenes are present rather than biotite and hornblende in the felsic gneisses, and the grade of metamorphism reaches granulite facies. The higher metamorphic grade may be a relict of the cooling history of the intrusive bodies.

Table 1 Brief outline of geological events recorded in the Kangerdlugssuaq region, eastern Greenland

1 Accumulation of interbedded volcanic and minor sedimentary units. Metamorphism and deformation produced interlayered mafic and felsic gneisses, strongly folded isoclinally.

2 *a*, Intrusion of granitic and trondhjemitic plutons accompanied, or closely followed by, strong deformation and recrystallisation which produced disharmonic folding of the interlayered mafic and felsic gneisses.
   b, Igneous activity was apparently protracted and in later stages produced mafic as well as potash-rich granitic plutons and pegmatites.

3 Subsequent to, and perhaps closely following, intrusive activity all the layered gneisses and plutonic bodies were folded and metamorphosed. This was followed, possibly at the end of folding, by pervasive shearing and by partial retrogressive recrystallisation.
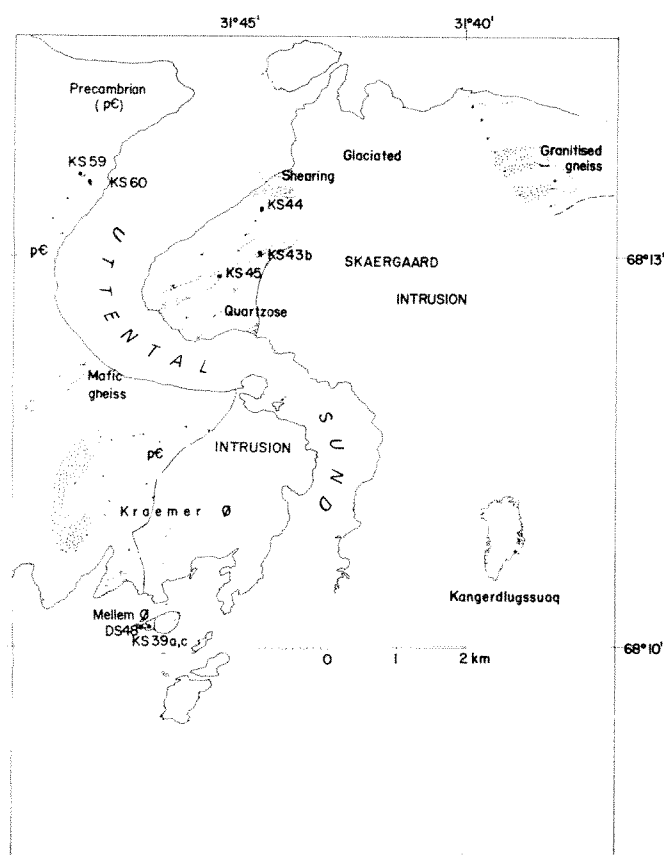
Fig. 1 Geological sketch map of the Kangerdlugssuaq area, showing locations of analysed samples.

Deformation of the gneisses is intense, with at least three major episodes of folding accompanying the several intrusive pulses. In the Kangerdluqssuaq area pervasive shearing (event 3, Table 1) has affected the continuity of earlier fold forms. Shearing was also accompanied by partial retrogressive recrystallisation. In areas of very intense shearing recrystallisation was more complete, pegmatitic gneisses were deformed to augen gneisses, and finer-grained rocks were mylonitised.

Lead isotope results

Samples from each of the gneiss sequences were selected for lead isotopic analysis; sample localities are shown in Fig. 1. Whole-rock powders were dissolved with HF and $HClO_4$, and lead was extracted using an HBr cation exchange procedure. Full procedural blanks amounted to 10–15 ng for 1-g samples (that is, less than 1% of the total quantity of Pb processed for any sample). Isotopic analyses were made on a 12-inch solid source mass spectrometer, using a silica gel–H_3PO_4 matrix. All analyses have been corrected for mass fractionation in the spectrometer by intercomparison with the NBS common lead standard (SRM-981). Corrected isotopic ratios are believed to be accurate to better than 0.15% of absolute values.

The isotopic analyses (Table 2) for the Kangerdlugssuaq gneisses are shown in a $^{207}Pb/^{204}Pb–^{206}Pb/^{204}Pb$ diagram in Fig. 2. If the inconsistent sample KS-60 (cataclasised augen gneiss) is excluded, the remaining seven samples lie within analytical uncertainty of an isochron with a slope of 0.220 ± 0.003 (1σ), corresponding to an age of $2,980 \pm 20$ (1σ) Myr BP (see ref. 11). Such a close fit to the Pb–Pb isochron suggests that these gneisses were derived from a source region that was homogeneous with respect to U/Pb during approximately single-stage evolution from the time of formation of the Earth to the isochron date. Geochemical differentiation at the latter time produced severe fractionation of U/Pb as well as isotopic homogenisation of Pb (that is, igneous parental material for

the gneisses acquired the mean isotopic composition of their source region at the isochron date, but had variable U/Pb). It is not clear whether isotopic rehomogenisation took place during the final metamorphism of gneisses emplaced over a significant period of time, or whether the gneisses were derived from the same homogeneous source over a relatively short period of time just before the culminating metamorphism. In the first case, the isochron date represents event 3 in Table 1, and a significant history of unknown duration is implied for the earlier events. In the second case, the isochron date may reflect the age of igneous material that was parental to the gneiss complex. But because the rocks associated with each sequence of events conform to the isochron, the latter possibility requires that the culminating metamorphism closely followed (within a few 10^7 yr) emplacement of the various parental rocks. We cannot yet resolve this problem, but because of the complex geological history recorded by the Kangerdlugssuaq rocks the first possibility seems more likely. Considering the wide spread in $^{206}Pb/^{204}Pb$ obtained for our samples, and the small residuals obtained in the isochron regression, it seems improbable that further whole-rock Pb isotope analyses of Kangerdlugssuaq gneisses will reveal ages older than about 3.0×10^9 yr.

Sample KS-60 may be inconsistent with the lead-isotope systematics of the other samples because of uranium loss (relative to lead) subsequent to intense shearing of this rock. But recent loss of uranium (for example during early Tertiary igneous activity or post-glacial uplift in the area) would not account for the comparatively low $^{207}Pb/^{204}Pb$ ratio in KS-60. It also seems unlikely that KS-60 was not initially affected by the event which caused effective homogenisation of lead isotopic compositions in the other gneiss samples (as seems required from the excellent fit of these samples to the $^{207}Pb/^{204}Pb–^{206}Pb/^{204}Pb$ isochron). Therefore, we think it likely that uranium loss occurred during, or shortly following, the shearing of KS-60.

Kangerdlugssuaq gneisses

A single-stage $^{238}U/^{204}Pb$ (μ_0) of 8.1 is indicated for the source materials of the Kangerdlugssuaq gneisses. This value of μ_0 corresponds to a lead isotope growth curve which intersects our $^{207}Pb/^{204}Pb–^{206}Pb/^{204}Pb$ isochron at 2,980 Myr and 0 Myr, based on values for primordial lead isotope ratios and decay constants given in ref. 11. This comparatively low value of μ_0 is compatible with a mantle or lower crust source region for igneous precursors to the gneisses, because upper crustal rocks are generally enriched in U/Pb. For example, such enrichment in U/Pb is inferred for the Isua supracrustal rocks which yield a μ_0 value of 8.7 (ref. 12). On the other hand, profound U depletion noted in many ancient infracrustal gneisses can

Table 2 Lead content and isotopic composition of Kangerdlugssuaq gneisses

| Field no.	p.p.m. Pb	$^{206}Pb/$ ^{204}Pb	$^{207}Pb/$ ^{204}Pb	$^{208}Pb/$ ^{204}Pb
KS-39A (1)*	9.19	16.601	15.299	37.568
Mafic gneiss with felsic layers				
KS-39C (2)*	13.31	17.164	15.369	35.908
Felsic gneiss				
KS-43B (3)*	4.30	14.861	14.886	34.587
Porphyritic felsic gneiss				
KS-44 (4)*	1.47	17.670	15.556	36.677
Amphibolite gneiss				
KS-45 (5)*	20.03	14.061	14.730	33.407
Pegmatitic plutonic gneiss				
KS-59 (6)*	30.00	20.434	16.169	33.876
Kf-rich augen gneiss				
KS-60 (7)*	12.67	20.467	15.841	60.468
Cataclasised augen gneiss				
DS-48 (8)*	39.83	23.387	16.749	36.760
Kf-rich granitic pegmatite				

*Numbers used in Fig. 2.
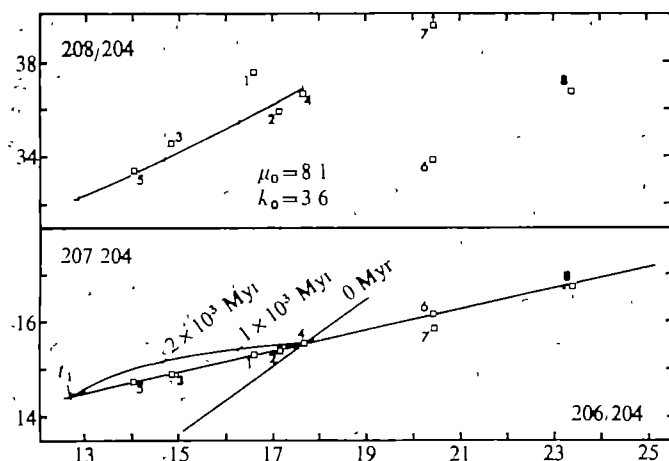
Fig. 2 Lead isotope systematics of analysed samples. A single-stage growth curve with a μ_o value of about 8.1 is shown for reference. t_1, 2,980 Myr.

result in present-day-unradiogenic Pb isotopic compositions. For comparison, 2.8×10^9 yr-old granulite facies rocks from south-western Greenland yield a μ_o value of 7.5 using the same decay constants[13]. The lower value of μ_o obtained for these granulite gneisses is in accord with the observation that U depletion commonly becomes more pronounced as metamorphic grade increases[14]. The 3.7×10^9 yr-old Amîtsoq gneisses of western Greenland also yield lower values of μ_o (about 7.6 to 7.7) and, in addition, contain Pb that is much less radiogenic ($^{206}Pb/^{204}Pb$ ranges from 11.5 to 14.9)[15,16] than the Kangerdlugssuaq gneisses. Consequently, it is unlikely that the Kangerdlugssuaq gneisses could represent metamorphosed parent material equivalent in composition to the Amîtsoq gneisses. Our work does not preclude the existence in the Kangerdlugssuaq area of older material with a composition different from that of the Amîtsoq gneisses. The Vikan migmatitic gneisses in northern Norway yield a Pb–Pb whole-rock isochron age of about 3.5×10^9 yr, and a value of μ_o of 8.9 (using the decay constants given in ref. 11)[17]. The range of $^{206}Pb/^{204}Pb$ values in these gneisses (14.8 to 21.2) is within the range for the Kangerdlugssuaq gneisses, but the latter display lower $^{207}Pb/^{204}Pb$ ratios in accordance with their younger age. Conceivably, high-grade metamorphism of Vikan-like rocks more than 3.0×10^9 yr ago could have pro-

duced gneisses with Pb isotopic compositions like those in the Kangerdlugssuaq rocks. A discussion of the origin of the Kangerdlugssuaq gneisses will be presented elsewhere.

The approximately 3.0×10^9-yr-old metamorphism recorded in the Kangerdlugssuaq gneisses was apparently related to regional metamorphism of high amphibolite or granulite grade (depending in part upon level of exposure) that affected much of the Greenland Shield between 3.0 and 2.7×10^9 yr ago[9,13,18–22]. High-grade metamorphic rocks of similar age have been documented from the Lewisian Complex of north-western Scotland[24–27]. It has been noted that this blanket metamorphism affected supracrustal as well as infracrustal rocks, thus recording an actual rise in crustal temperature during that time[9]. The complex geological relationships recorded in the Kangerdlugssuaq gneisses, like those occurring in the Lewisian gneisses[10], suggest a more extensive history of crustal evolution that predates the $3.0–2.7 \times 10^9$-yr-old event. U–Pb zircon and Rb–Sr whole-rock analyses are in progress in an attempt to resolve earlier events and to test our interpretation that the whole-rock Pb–Pb isochron age for the Kangerdlugssuaq gneisses reflects the time of last major metamorphism in that area.

1 Wager, L. B., Meddr. Grønland, 105, 1 (1934).
2 Bertholsson, A., and Noe-Nygaard, A., The Precambrian, 2 (edit. by Rankama, K.), 113 (1965).
3 Wager, L. R., and Hamilton, E. I., Nature, 204, 1079 (1964).
4 Turner, G., Palaeogeophysics (edit by Runcorn, S K), 491 (Academic, London, 1970).
5 Wager, L. R., and Deer, W. A , Meddr. Grønland, 105 (4), 1 (1939).
6 Bockinsale, R. D., Brooks, C K , and Rex, D. C , Bull. geol Soc. Denm., 20, 27 (1970).
7 Pankhurst, R. J., Bockinsale, R D , and Brooks, C. K , Contr. Miner Petrol., 54, 17 (1976).
8 McGregor, V. R., Phil. Trans. R Soc., A273, 343 (1973).
9 Bridgwater, D., Watson, J., and Windley, B F , ibid , 273, 493 (1973).
10 Davies, F. B , Nature, 258, 589 (1975).
11 Tatsumoto, M , Knight, R J , and Allegre, C. J , Science, 180, 1279 (1973).
12 Moorbath, S , O'Nions, R K , and Pankhurst, R. J , Nature, 245, 138 (1973).
13 Black, L P., Gale, N H , Moorbath, S , Pankhurst, R J , and Windley, B. F , Nature, 244, 50 (1973).
14 Heier, K S , Phil. Trans. R. Soc., A273, 429 (1973).
15 Black, L P , Gale, N. H , Moorbath, S., Pankhurst, R J , and McGregor, V. R , Earth planet. Sci. Lett., 12, 245 (1971).
16 Moorbath, S , O'Nions, R K., and Pankhurst, R J , ibid , 27, 229 (1975).
17 Taylor, P N., ibid , 27, 35 (1975).
18 Bridgwater, D , Rapp Grøn geol Unders , 55, 51 (1973).
19 Pankhurst, R J., Moorbath, S , and McGregor, V. R , Nature, 243, 24 (1973).
20 Rex, D. C., and Gledhill, A , Bull geol Soc Denmark, 23, 29 (1974).
21 Pidgeon, R T., and Hopgood, A M , Geochim cosmochim. Acta, 39, 1333 (1975).
22 Nunes, P D , Stenger, R H , and Bridgwater, D , Rapp. Grøn geol Unders , 66, 21 (1974).
23 Gulson, B. L., and Krogh, T E., Geochim. cosmochim. Acta, 39, 65 (1975).
24 Moorbath, S., Welke, H., and Gale, N H , Earth planet Sci. Lett , 6, 245 (1969).
25 Pidgeon, R T , and Bowes, D R , Geol. Mag , 109, 247 (1972).
26 Moorbath, S , Powell, J. L., and Taylor, P. N , J geol Soc Lond , 131, 213 (1975).
27 Bikerman, M , Bowes, D. R., and van Broomen, O , ibid , 131, 237 (1975).

Abnormal or absent β mRNA in β⁰ Ferrara and gene deletion in δβ thalassaemia

Francesco Ramirez, Joyce V. O'Donnell, Paul A. Marks & Arthur Bank
Columbia University, College of Physicians and Surgeons, New York

Salvatore Musumeci, Gino Schilirò, Giuseppe Pizzarelli & Giuseppe Russo
Pediatric Clinic, University of Catania, Italy

Benedetta Luppis & Roberto Gambino
Laboratory of Molecular Embryology, Arco Felice, Naples, Italy

In patients with β⁰ thalassaemia from Ferrara, β globin mRNA sequences are either absent or structurally abnormal while in β⁰ thalassaemia in Catania, β globin mRNA sequences are present. In δβ thalassaemia there is a deletion of β-like globin genes, while in β⁰ Catania DNA, no β globin gene deletion is detectable.

THREE general types of β thalassaemia have been described to date: β⁺ thalassaemia, associated with the production of decreased amounts of structurally normal β globin; β⁰ thalassaemia, characterised by the absence of β globin synthesis; and δβ thalassaemia, in which synthesis of both δ and β globin is absent[1]. The defect in β globin synthesis in these syndromes can be demonstrated by measurements of globin synthesis both in whole cells and by addition of isolated mRNA to
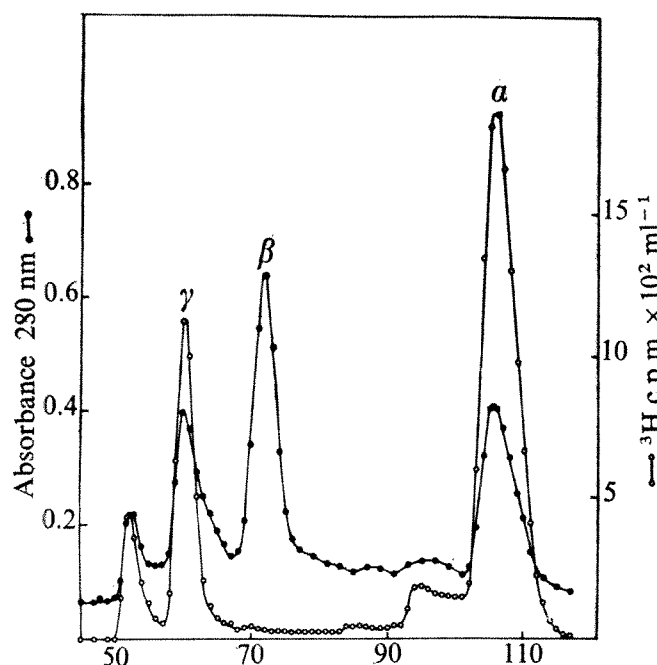
Fig. 1 Globin synthesis in reticulocytes of β^0 Ferrara. Elution from carboxymethyl cellulose chromatography of globin prepared from cells of a patient homozygous for β^0 Ferrara (patient 2 in Fig. 4b and Table 1) which had been incubated with ^{3}H-leucine.[4] The patient had been transfused 22 d before the blood sample was obtained.

heterologous cell-free systems[2-5]. More recently, the availability of complementary DNAs (cDNAs) specific for α and β globin mRNA sequences has made it possible to quantify the α and β mRNA content of cells in these disorders[6-12]. In the β^+ thalassaemias, we and others[6,7] have shown previously, using these molecular probes, that the amounts of cytoplasmic β globin mRNA is decreased. By contrast, in β^0 thalassaemia in patients from Italy, there have been conflicting reports[8,11] either of the absence of structurally normal β mRNA sequences, or of structurally abnormal β mRNA sequences. β mRNA sequences in $\delta\beta$ thalassaemia have been shown previously to be absent[8]. We report here that in β^0 thalassaemia in patients from the Ferrara region of Italy, another distinctive genetic defect may be reflected at the mRNA level by the inability of β^0 Ferrara mRNA to hybridise fully to β cDNA. In contrast, we find significant amounts of β mRNA in β^0 Catania mRNA, which is consistent with the presence of abnormal, untranslated mRNA in some β^0 thalassaemia patients[6,11]. We also report that

Table 1 Amounts of α and β mRNA and percentage β cDNA hybridised

Source of RNA*	α/β mRNA content†	% β cDNA hybridised‡
Normal (8)§	0.9–1.36	90–96
β^+ thalassaemia (6)	4.5–14.3	82–88
$\delta\beta$ thal homozygote	0	18
β^0 Catania —1‖	2	80
—1‖	2.4	80
β^0 Ferrara —1	1.8	50
—2	1.78	54
—3	1.43	46
—4	1.04	56
—5	0.4	48

*6 to 16S RNA was isolated by sucrose density gradient centrifugation in all cases from peripheral blood.
†$C_0t_{\frac12}$ β cDNA/$C_0t_{\frac12}$ α cDNA. The $C_0t_{\frac12}$ is the C_0t at the midpoint between the initial and final hybridisation plateau.
‡The % of β cDNA resistant to micrococcal nuclease digestion at the final hybridisation plateau. The background without added RNA has not been subtracted.
§Number of patients studied is in parentheses.
‖Studies performed twice on the same patient.

in $\delta\beta$ thalassaemia there is a deletion of structural δ and/or β globin gene sequences; in β^+ and β^0 Catania DNA, no such deletion is detectable. These results indicate a heterogeneity of gene defects in β thalassaemia, all of which can lead to decreased or absent β globin synthesis.

The β^0 Catania patients are homozygous for β^0 thalassaemia and manifest the clinical symptoms of β thalassaemia. They range in age from 5 to 10. All of the patients have hepatosplenomegaly, abnormal red cell morphology with hypochromic, microcytic anaemia, elevated haemoglobin A_2 and F levels and require blood transfusions. The $\delta\beta$ thalassaemia homozygote is a patient who has only haemoglobin F, a mild anaemia, and thalassemic smear with hypochromic, microcytic cells[13]. The β^0 Ferrara patients have high A_2-type homozygous β thalassaemia[14]. Fibroblasts from patients with hereditary persistence of foetal haemoglobin (HPFH) and from normal patients were cultured by Dr Robert Krooth. The HPFH fibroblasts are from a patient who is homozygous for the black form of this disorder[15].

Total RNA was extracted from peripheral blood samples from which buffy coat had been removed[16]. RNA was fractionated on sucrose density gradients and the 6–16S region was used in hybridisation and biological activity studies. α and β globin cDNAs were purified as previously described[9]. The α and β globin cDNAs sediment as a relatively homogeneous 7S peak using alkaline sucrose gradient analysis. DNA was prepared essentially as described previously[17]. Hybridisation of globin cDNA with spleen, liver, fibroblast and white blood cell DNA was also performed as previously described[9].
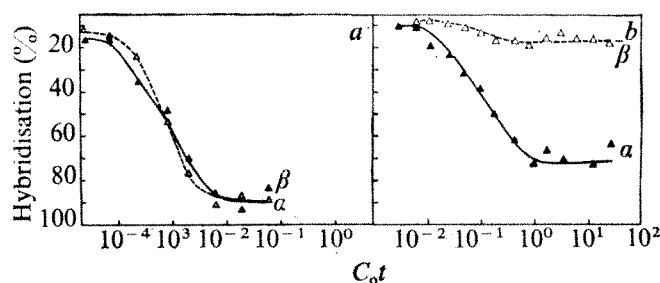
Fig. 2 Hybridisation of α and β cDNA with normal (a) and homozygous $\delta\beta$ thalassaemia RNA (b). Between 2,000 and 2,500 c.p.m. of α or β cDNA were hybridised to 5×10^{-6} to 1.2×10^{-2} µg of normal 6 to 16S RNA for 4 h, and to 1×10^{-4} to 1 µg of $\delta\beta$ thalassaemia 6 to 16S RNA for 24 h to obtain the C_0t values shown. Hybridisation was performed in a 10-µl reaction mixture at pH 7.0 containing (final concentration): 0.02 M sodium phosphate, 0.1% sodium dodecyl sulphate, 0.3 M NaCl and 0.002 M EDTA at 68 °C. The percentage hybridisation was determined after digestion of single-stranded cDNA with micrococcal nuclease: 4-µl samples of each hybridisation mixture were treated with micrococcal nuclease and 4 µl was used as control for recovery in individual experiments. ▲, α cDNA; △, β cDNA.

No β globin synthesis was demonstrable by column chromatographic analysis when peripheral blood samples containing reticulocytes from β^0 Catania, β^0 Ferrara or $\delta\beta$ thalassaemia samples were incubated with ^{3}H-leucine. A typical chromatographic separation on carboxymethyl cellulose[4] in a patient from Ferrara is shown in Fig. 1. In contrast to previous studies[18], there is no induction of β globin synthesis 22 d after blood transfusion. Similar results were obtained with β^0 Catania cells and with homozygous $\delta\beta$ thalassaemia cells[13]. When mRNAs from $\delta\beta$, β^0 Catania and β^0 Ferrara patients are added to a Krebs' ascites tumour cell-free system (ref. 19 and R. G. and B. L., unpublished) there is also no detectable synthesis of β globin chains. Only α and γ globin synthesis are present.

Using normal reticulocyte 6 to 16S RNA, the relative content of α and β mRNA measured by hybridisation to purified cDNAs is close to 1.0 (Table 1, Fig. 2a). In addition, greater than 90% of the α and β cDNAs are protected from micro-

coccal nuclease digestion by normal reticulocyte mRNA, indicating that all of the nucleotide sequences in the cDNA are represented in the mRNA. Using 6–16S RNA isolated from the δβ thalassaemia homozygote, the α cDNA hybridises to a saturation plateau of 70% at a relatively low C_0t value while at a C_0t value 30–40-fold above this, only 18% of the β cDNA is hybridised (Table 1, Fig. 2b). The background of the hybridisation in the absence of added mRNA is 9%. These data indicate that δβ thalassaemia homozygote mRNA contains no detectable δ or β mRNA sequences, and confirms previous studies[8]. In addition, they show that the β cDNA probe used is contaminated with less than 10% α cDNA. The fact that only 70% of the α cDNA is hybridised in a vast excess of δβ thalassaemia mRNA (Fig. 2b) suggests that the α cDNA is contaminated to approximately 20–25% with β cDNA sequences. Additional evidence supporting this is obtained when the α cDNA is hybridised with hydrops fetalis mRNA (containing no α mRNA)[9]; only 20 to 25% of the α cDNA is protected while β cDNA is hybridised to 85% (Fig. 3a).
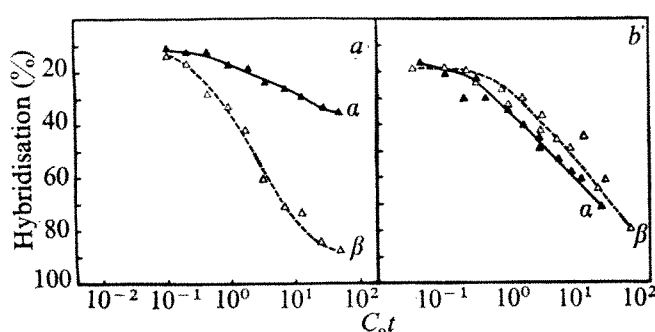
Fig. 3 Hybridisation of α and β cDNA with hydrops fetalis (a) and β⁰ Catania RNA (b). 2,000 and 2,500 c.p.m. of α or β cDNA were hybridised to 2×10^{-2} to 10 μg of hydrops fetalis 6 to 16S RNA for 4 h and to 1.3×10^{-4}–6.4×10^{-1} μg of β⁰ Catania 6–16S RNA for 120 h. The hybridisation conditions and analysis are as described in the legend to Fig. 2. ▲, α cDNA; △, β cDNA.

When 6–16S mRNA from β⁰ Catania cells is hybridised to α and β cDNAs, a completely different hybridisation curve is obtained. The $C_0t_{\frac{1}{2}}$ with the β cDNA is 2-and 2.4-fold greater than that with α cDNA indicating 2-to 2.4-fold greater amounts of α than β globin mRNA sequences are present (Fig. 3b, Table 1). Here, 80% of the β cDNA is protected from micrococcal nuclease digestion using the RNA of this patient. Although it is clear from Fig. 3b that almost all of the β cDNA is hybridised and the curve is continuing to descend, a high enough C_0t value was not attained to determine if all of the nucleotide sequences in the β cDNA are present in the β mRNA molecules in the cells of this patient. More extensive studies at higher C_0t values in this and other patients with β⁰ Catania thalassaemia are required to determine if all of the sequences present in intact β mRNA are present. However, the extent of β cDNA hybridisation is clearly different from that of δβ thalassaemia mRNA. These data are consistent with those reported previously in two patients using rabbit globin cDNA probes[6] and in two Chinese patients and a β⁰ patient of Italian extraction[11] suggesting the possibility that there is non-functional β globin mRNA in some patients with β⁰ thalassaemia. Since in these[6,11] and the present studies, less than full-length β cDNA probes are used, even complete hybridisation of the β cDNA would not prove that complete β globin mRNA molecules are present and untranslatable. To do this, studies using full-length cDNA and structural analysis of the β mRNA will be required. The results reported here also do not exclude the possibility that some patients with β⁰ thalassaemia from Catania may have

an absence of β mRNA as has been reported in other β⁰ patients[8,12].

Defect in mRNA from β⁰ Ferrara cells

When β⁰ Ferrara 6–16S RNA is hybridised to the α and β cDNA probes, a different result is obtained from that with normal, δβ or β⁰ Catania mRNA. First, the β cDNA is only hybridised to 46–56% at C_0t values even in mRNA excess 10- to 25-fold greater than that necessary to protect the α cDNA maximally (Table 1, Fig. 4). Using cDNA synthesised from HbH mRNA, similar extents of hybridisation were obtained to those using β cDNA, 45–55%. The $C_0t_{\frac{1}{2}}$ values are an accurate measure of the relative amounts of α and β mRNA present in a sample of RNA[6,9,20]. The $C_0t_{\frac{1}{2}}$ values for α in Fig. 4 are 2×10^{-2} and 9×10^{-2} respectively. If a small amount of intact β mRNA capable of hybridising to all of the β cDNA is present in β⁰ Ferrara mRNA, it must represent less than 5% and 2% of α mRNA since the $C_0t_{\frac{1}{2}}$ of any complete β mRNA must be at least 4×10^{-1} and 6 (highest C_0t values reached) respectively.

Mixing of two β⁰ Ferrara mRNA preparations in different combinations did not increase the percentage hybridisation of β cDNA. By contrast, mixing of β⁰ Ferrara mRNA with hydrops fetalis mRNA or normal mRNA leads to over 90% hybridisation of the β cDNA, indicating that no inhibitor of β mRNA hybridisation is present in the β⁰ Ferrara mRNA samples. When β⁰ Ferrara is hybridised to β cDNA at 78 °C instead of 68 °C, there is a decrease in the amount of cDNA protected from micrococcal nuclease digestion to a background value. The hybridisation of normal reticulocyte mRNA to β cDNA and of normal and β⁰ Ferrara mRNA to α cDNA are unaffected by a similar increase in temperature. The lower level of protection of α cDNA by β⁰ Ferrara mRNA than by normal mRNA can be explained by contamination of α cDNA by β cDNA sequences. This was confirmed by showing that when hydrops fetalis mRNA containing only β and γ mRNA was added to the Ferrara mRNA, 90% of the α cDNA probe could be protected.

In β⁰ Ferrara mRNA, it is clear that different hybridisation results are obtained than with β⁺ and β⁰ Catania and δβ mRNA. Even in large mRNA excess only 46–56% of the β cDNA is hybridised; only about half of the nucleotide sequences present in β cDNA are represented in β Ferrara RNA in significant amounts. This finding cannot be explained by contamination of the β cDNA probe by α cDNA since only 18% of the β cDNA is hybridised at similar inputs of RNA from δβ thalassaemia cells which also contain α mRNA in normal amounts. In all five β⁰ Ferrara patients studied, the defects in β mRNA are similar if not identical since there is no increase in saturation

Fig. 4 Hybridisation of α and β cDNA with β⁰ Ferrara–1 RNA (a) and β⁰ Ferrara–2 RNA (b). Between 2,000 and 2,500 c.p.m. of α or β cDNA were hybridised to 4×10^{-4} μg to 3×10^{-1} μg of β⁰ Ferrara–2 RNA for 8 h, and to 1.2×10^{-4}–8×10^{-2} μg β⁰ Ferrara–1 RNA for 4 h. The hybridisation conditions and analysis were as described in the legend to Fig. 2. ▲, α cDNA; △, β cDNA.
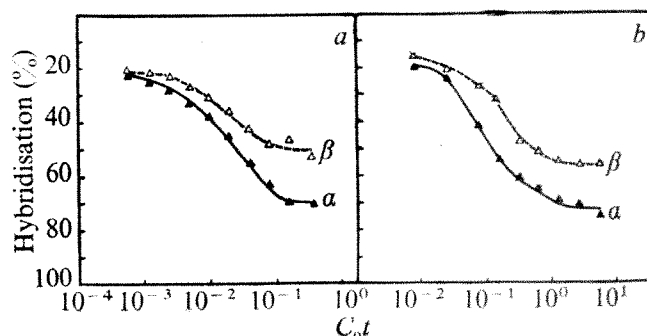
hybridisation with β cDNA when mixing of β⁰ Ferrara mRNA samples were performed.

The possibility that the β cDNA is hybridising with γ mRNA sequences in β⁰ mRNA was studied by comparing the hybridisation of γ cDNA prepared as described previously[21] with δβ thalassaemia and β⁰ Ferrara mRNA. Using both types of mRNA the γ cDNA probe was hybridised at $C_0t_{1/2}$ values two-to threefold less than using α cDNA, indicating approximately one-third to one-half as much γ as α mRNA is present in both of these mRNAs (F.R., J.V.O'D., C. Natta and A.B., unpublished). The hybridisation data support cell-free studies indicating that similar amounts of γ globin mRNA are present in β⁰ Ferrara and δβ mRNA. The results with δβ mRNA indicate that γ mRNA does not cross hybridise with the β cDNA probe in the conditions of hybridisation, and that the difference in the extent of hybridisation of β cDNA to δβ and β⁰ Ferrara mRNA cannot be explained by the presence of γ mRNA.

Several possible mechanisms can explain the 35–45% difference in hybridisation of β⁰ Ferrara mRNA and δβ thalassaemia mRNA with β cDNA. First, β mRNA may be completely absent in these cells, but since δ mRNA is present in β⁰ Ferrara cells, the partial hybridisation may be due to homology between β cDNA and δ mRNA nucleotide sequences. Since β and δ globin chains differ by only 9 of 144 amino acids, significant homology at the nucleotide level is not unlikely. But the relatively low C_0t of hybridisation to β cDNA when compared with hybridisation to α cDNA (Fig. 4) requires a rather large amount of δ mRNA to be present in these cells.

If incomplete or structurally abnormal mRNA existed, the abnormality in β⁰ Ferrara mRNA could be either at the 5' or 3' end of the mRNA outside of the nucleotide sequence necessary to encode β globin; the complete nucleotide sequence in mRNA necessary to encode β globin could be intact, although this mRNA might be more labile than normal mRNA and reports of induction of β globin synthesis by transfusion *in vivo* and addition of supernatant fractions *in vitro* would be possible[18,23]. Regardless of the precise explanation for the limited hybridisation of Ferrara mRNA to β cDNA, the results suggest that the genetic lesion in β⁰ Ferrara is different from that in β⁰ Catania, β⁺ and δβ thalassaemia and is associated with either an absence of β globin mRNA or an abnormal β globin mRNA. The precise nature of the gene defects in the β⁰ Catania and β⁰ Ferrara and perhaps other forms of β⁰ thalassaemia must await analysis of newly synthesised β globin mRNA sequences in the nuclei of erythroid cells in patients with these syndromes.

Deletion of globin gene sequences in δβ thalassaemia

When β cDNA is hybridised to normal DNA approximately 51–56% of the β cDNA is hybridised at similar cDNA : cellular DNA inputs (Table 2). DNA from normal fibroblasts, spleen and white blood cells all give similar levels of hybridisation. When DNA from β⁰ Catania cells is hybridised to β cDNA, approximately 50–55% of the β cDNA becomes hybridised. These results indicate that there is no detectable deletion of

Table 2 Hybridisation of α and β cDNA to cellular DNA

| | Percent hybridisation* | | | |
| | α cDNA | | β cDNA | |
Source of DNA	C_0t : 4,320	C_0t : 8,640	C_0t : 4,320	C_0t : 8,640
Normal – 1	43	39	51	51
Normal – 2	47	49	54	56
Normal – 3	47	46	51	51
HPFH homozygote	35	36	23	22
δβ homozygote	37	41	32	29
δβ heterozygote – 1	44	44	38	41
δβ heterozygote – 2	40	43	41	43
β⁰ Catania – 1	47	46	55	50
β⁰ Catania – 2	45	43	56	54
Hydrops fetalis	21	20	53	53

*In each cDNA–DNA hybridisation, 3.3 × 10⁻⁷ as much cDNA as cellular DNA is used. Forty-five μl aliquots containing 0.135 mg of DNA and 630 c.p.m. (0.045 ng) of either α or β cDNA were mixed in 0.12 M sodium phosphate (pH 6.8) and 0.4% sodium dodecyl sulphate in 100-μl capillary pipettes, heated in a boiling water bath for 15 min and incubated at 68 °C for varying times in order to attain the C_0t values shown. The cDNA eluted from hydroxylapatite at 68 °C by 0.4 M sodium phosphate is expressed as a percentage of cDNA eluted by 0.12 M plus 0.4 M sodium phosphate successively. The values obtained at different C_0t values are shown to indicate a saturation plateau has been reached.

Alternatively, the RNA sequences hybridising to β cDNA may represent Lepore-like globin mRNA in the Ferrara RNA, and the hybridisation to β cDNA due to homology with the β-like sequences in this Lepore-like chain. Or third, there may be a partial deletion of β globin mRNA sequences at the 3' end of Ferrara mRNA. In addition to 50–100 polyadenylate residues at the 3' end of human globin mRNA[22], studies of abnormal elongated human globin mutants suggest that there is an untranslated nucleotide sequence of approximately 100 nucleotides at this end[1]. Since β globin cDNA contains approximately 400 nucleotides and primarily sequences synthesised from the 3' end of mRNA, a deletion at the 3' end of the Ferrara mRNA might result in the findings reported here. Lastly, the presence of normal amounts of a structurally abnormal mRNA with either increased lability or considerable secondary structure must be considered. Such an abnormal RNA might not be fully hybridised to β cDNA. Either cross hybridisation of β cDNA to δ mRNA sequences or an abnormal β mRNA could explain the lability of the β⁰ Ferrara mRNA–β cDNA hybrid.

β or β-like genes in β⁰ thalassaemia. In contrast, when HPFH DNA is mixed with β cDNA only 22–23% of the β cDNA is hybridised at saturation (Table 2). Using DNA from a δβ thalassaemia homozygote, only 32 and 29% of the β cDNA is hybridised, somewhat higher than that using HPFH DNA (Table 2). Similarly, DNA from heterozygotes for δβ thalassaemia show decreased hybridisation to the β cDNA probe to 41 and 43%, respectively (Table 2). To determine if the deletions in HPFH and δβ thalassaemia DNA are overlapping, 0.112 mg of homozygous δβ thalassaemia DNA and 0.113 mg of HPFH DNA were mixed and annealed with 0.045 ng β cDNA to a C_0t of 8,640; 37% of the β cDNA was protected. 0.225 mg of normal DNA gave 59% hybridisation while the same amount of HPFH DNA gave 24% hybridisation with the same input of β cDNA. The data are consistent with overlapping gene deletions in HPFH and δβ thalassaemia DNA with a larger deletion in HPFH DNA.

When α cDNA is hybridised to cellular DNA from normal, HPFH, δβ and β⁰ thalassaemia cells at similar inputs of cDNA

and cellular DNA, relatively similar amounts of hybridisation of the α cDNA are obtained (Table 2). These values indicate that normal, HPFH, and β^0 and $\delta\beta$ thalassaemia DNA all contain similar numbers of α globin genes and serve as a useful internal control of the cellular DNA : cDNA input. There is 5 –10% less hybridisation of α cDNA with HPFH, and 3–7% less hybridisation of α cDNA with $\delta\beta$ thalassaemia DNA from the homozygote than with similar inputs of normal or β^0 thalassaemia DNA (Table 2). This is consistent with β cDNA sequences in the α cDNA probe which cannot be hybridised because of the deletion of β-like cellular DNA sequences. The somewhat higher level of hybridisation of normal DNA to cDNA than to α cDNA (51–56% compared to 39–49%) confirms previous data[9].

Studies to date indicate that in homozygous β^+ thalassaemia[9], β^0 thalassaemia[12] and in a $\delta\beta/\beta^0$ double heterozygote[10] there is no detectable deletion of globin structural genes. In contrast, deletion of globin structural genes has been reported in HPFH homozygotes[15,24]. Our results confirm deletion of β-like globin DNA nucleotide sequences in HPFH DNA (Table 2). Less than half as much hybridisation of β cDNA occurs with HPFH DNA as with similar input of normal DNA. The nature of the sequences in HPFH DNA which hybridise to β cDNA is not yet defined; they may represent incompletely deleted δ and/or β structural gene loci, embryonic β-like genes or cross hybridisation of α gene sequences to the α cDNA contaminating the β cDNA probe.

Deletion of β structural gene sequences has recently been observed by others (S. Ottolenghi *et al.*, unpublished) in patients homozygous for $\delta\beta$ thalassaemia. We have confirmed the deletion of β-like gene material in one of these homozygotes. Further, we demonstrate detectable deletions of β-like gene material in heterozygotes for $\delta\beta$ thalassaemia (Table 2). Our methodology seems to be more sensitive in detecting gene deletions in $\delta\beta$ heterozygotes than hybridisation in vast DNA excess or vast cDNA excess, since such a deletion was not detected by others using these methods[10]. These data are consistent with at least partial deletion of β or δ globin structural genes or parts of both in this disorder. The deletion of structural globin gene sequences in $\delta\beta$ thalassaemia DNA may, in addition, be less than that in HPFH DNA. In both HPFH and homozygous $\delta\beta$ thalassaemia, the only haemoglobin produced is foetal haemoglobin, HbF. In HPFH, sufficient γ globin is produced to maintain normal haemoglobin levels, while in $\delta\beta$ thalassaemia, γ globin production does not compensate and there are cellular abnormalities typical of thalassaemia and consequent haemolytic anaemia. Regulation of γ globin synthesis at the gene level has been postulated to involve regulatory nucleotide sequences adjacent to the δ structural genes[25]. It may be speculated that deletion of these sequences in HPFH permits maximal expression of γ globin genes even in adult life. The differences in hybridisation between HPFH and $\delta\beta$ thalassaemia DNA with β cDNA reported here could reflect the presence of residual sequences in $\delta\beta$ thalassaemia DNA with a suppressive effect on γ globin production. The residual sequences present in $\delta\beta$ thalassaemia and absent in HPFH DNA may generate a gene product which suppresses γ gene activity. Alternatively, no such diffusible gene product acting in *cis* would be required if the sequences we are detecting in $\delta\beta$ thalassaemia which are absent in HPFH are adjacent to the γ globin genes and are involved in the normal suppression of γ globin gene activity in adult life. These speculations do not exclude the possibility that other genetic interactions may control γ gene expression.

Recently, a patient with β^0 thalassaemia has been described in whom there is absence of β mRNA but no apparent deletion of β-like genes[12]. In the present and previous studies[9,12] of patients with β^0 and β^+ thalassaemia who have intact δ globin loci, it is possible that cross hybridisation between β cDNA and δ gene sequences can obscure the detection of small deletions in β structural genes. In addition, the use of less than complete length β cDNAs in all of the studies reported to date

of gene deletion in β and $\delta\beta$ thalassaemia make it possible that deletions of β-like sequences in DNA at the end of β or δ genes homologous to the 5' end of β globin mRNA may exist and remain undetected in these syndromes.

The common denominator of the β thalassaemia syndromes is decreased β globin synthesis with inadequate γ globin compensation. The decrease in β globin production in β^+ thalassaemia is associated with decreased β globin mRNA and is due to either a defect in regulatory genes controlling β structural gene expression or the production of a structurally abnormal β gene product which is partially destroyed during RNA processing in the nucleus. In one patient with β^0 thalassaemia, absent β mRNA is associated with no detectable deletion of β globin genes[12]. On the other hand, as demonstrated in this paper, the absent δ and β mRNA in $\delta\beta$ thalassaemia is associated with deletion of δ and/or β gene material. In some β^0 thalassaemia patients[6,11] including one from Catania described here there are significant amounts of incomplete or complete β mRNA sequences in cells which are untranslated and may reflect a small abnormal nucleotide sequence in the β globin gene. This report provides evidence that β^0 Ferrara may differ from other forms of β^0 thalassaemia described to date and is associated with either abnormal or absent β mRNA. Thus, the β thalassaemia syndromes represent a heterogeneous group of disorders at the genetic level, all of which lead to decreased β globin production.

1 Weatherall, D. J., and Clegg, J. B., *The Thalassemia Syndromes*, 2nd ed. (Blackwell, Oxford, 1973).
2 Nienhuis, A. W., and Anderson, W. F., *J. clin. Invest.*, 50, 2458–2460 (1971).
3 Benz, E. J., and Forget, B. G., *J. clin. Invest.*, 50, 2755–2760 (1971).
4 Dow, L. W., *et al.*, *Nature new Biol.*, 243, 114–116 (1973).
5 Benz, E. J., Swerdlow, P. S., and Forget, B. G., *Blood*, 45, 1–10 (1975).
6 Kacian, D. L., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 70, 1886–1890 (1973).
7 Housman, D., Forget, B. G., Skoultchi, A., and Benz, E. J., Jr, *Proc. natn. Acad. Sci. U.S.A.*, 70, 1809–1813 (1973).
8 Forget, B. G., Benz, E. J., Jr, Skoultchi, A., Baglioni, C., and Housman, D., *Nature*, 247, 379–381 (1974).
9 Ramirez, F., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 72, 1550–1554 (1975).
10 Ottolenghi, S., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 72, 2294–2299 (1975).
11 Kan, Y. W., Holland, J. P., Dozy, A. M., and Varmus, H. E., *Proc. natn. Acad. Sci. U.S.A.*, 72, 5140–5144 (1975).
12 Tolstoshev, P., *et al.*, *Nature*, 259, 95–98 (1976).
13 Russo, G., Musumeci, S., Schilirò, G., D'Agata, A., and Pizzarelli, G., *Atti della. Società Italiana di Ematologia*, 1, 197–205 (1974).
14 Bargellesi, A., Pontremoli, S., and Conconi, F., *Eur. J. Biochem.*, 1, 71–73 (1937).
15 Kan, Y. W., Holland, J. P., Dozy, A. M., Charache, S., and Kazazian, H. H., *Nature*, 258, 162–163 (1975).
16 Terada, M., *et al.*, *Biochem. biophys. Res. Commun.*, 47, 766–774 (1972).
17 Gambino, R., *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, 71, 3966–3970 (1974).
18 Conconi, F., *et al.*, *Nature*, 254, 256–259 (1975).
19 Metafora, S., Terada, M., Dow, L. W., Marks, P. A., and Bank, A., *Proc. natn. Acad. Sci. U.S.A.*, 69, 1299–1303 (1972).
20 Ramirez, F., *et al.*, *J. biol. Chem.*, 250, 6054–6058 (1975).
21 Lanyon, W. G., Ottolenghi, S., and Williamson, R., *Proc. natn. Acad. Sci. U.S.A.*, 72, 258–262 (1975).
22 Cann, A., Gambino, R., Banks, J., and Bank, A., *J. biol. Chem.*, 249, 7536–7540 (1974).
23 Conconi, F., Rowley, P. T., del Senno, L., Pontremoli, S., and Volpato, S., *Nature*, 238, 83–85 (1972).
24 Forget, G. J., *et al.*, *Cell*, 7, 323–329 (1976).
25 Huisman, T. H. J., *et al.*, *Ann. N.Y. Acad. Sci.*, 232, 107–124 (1974).

Enzyme evolution in a microbial community growing on the herbicide Dalapon

E. Senior & A. T. Bull*

Biological Laboratory, University of Kent at Canterbury, Canterbury CT2 7NJ, UK

J. H. Slater

Department of Environmental Sciences, University of Warwick, Coventry CV4 7AL, UK

A seven-membered microbial community capable of utilising the herbicide Dalapon has been isolated by continuous-flow enrichment culture. The composition of this community has remained remarkably stable over thousands of hours in a Dalapon-limited chemostat. During this period, however, one member of the community, Pseudomonas putida, acquired the ability to grow on Dalapon through the evolution of an extant dehalogenase.

IN most natural environments different populations of microorganisms grow in close proximity to each other and may utilise similar substrates present in the ecosystem. One likely consequence of this is that various nutritional and physical interactions will develop between the populations leading to the evolution of complex, mutually dependent and possibly highly stable microbial communities. It is reasonable to postulate that interacting microbial communities are common in nature; however, very few microbial associations of this type have been defined, principally because inappropriate enrichment and selection techniques have been used to isolate them from natural environments.

The aims of the work reported here was to use suitable continuous-flow culture enrichment procedures[1,2] to isolate a microbial community which could grow on the herbicide 2,2'-dichloropropionic acid (Dalapon, Dow Chemical Company) as its sole source of carbon and energy. Dalapon is a widely used herbicide which is effective against monocotyledons and is known to undergo microbial degradation[3]. Such a compound was selected because it was argued that mixed populations possessing diverse metabolic capabilities are more likely to be able to metabolise recalcitrant or semi-recalcitrant compounds than are populations composed of a single species. In addition, it seems to be the general case that so called 'environmentally foreign' compounds are degraded much more rapidly in natural environments than they are by monocultures of microorganisms grown in similar conditions in the laboratory. Such observations provide circumstantial evidence for cooperative metabolic attack by a microbial community.

Characterisation of isolated microbial community

Figure 1 describes the structure of the microbial community which has been isolated on several occasions from four different soil sources, two of which had been pretreated with Dalapon. Initially, after three weeks growth in the chemostat enrichment system, the mixed culture consisted of seven different microorganisms which could be divided into two groups. The first group comprised the primary utilisers which, in monoculture, were able to grow on Dalapon as the only carbon and energy source. In all four microbial communities originally isolated, three primary utilisers were

found; two were Gram-negative, motile rod bacteria, one of which (P1) was identified as a *Pseudomonas* species of Shewan's group II (ref. 4) the other (P2) remains unidentified. The third Dalapon utiliser was a filamentous fungus, *Trichoderma viride* (P4). The other group contained the secondary utilisers which were unable to grow on Dalapon in monoculture in any growth conditions. These secondary organisms must have been growing on metabolites which were produced directly from the primary catabolism of Dalapon, or metabolites excreted by the growing primary utilisers, or on compounds freed into the growth medium as a consequence of the death and lysis of primary and secondary utilisers. The non-Dalapon utilisers were three bacteria —a *Flavobacterium* species (S1), *Pseudomonas putida* (S3), and an unidentified pseudomonad (S4)—and a pink budding yeast (S5). The quantitative species composition remained very stable with changing growth conditions with the exception of the two major primary Dalapon utilisers; at low dilution rates in the chemostat (less than $0.25\ h^{-1}$) the bacterium P2 was the dominant primary organism whereas at high dilution rates (greater than $0.45\ h^{-1}$) *Pseudomonas* species P1 dominated.

Stability of the community

The yeast was only loosely associated with the microbial community and was eliminated from the enrichment chemostat the first time the dilution rate was raised above $0.2\ h^{-1}$.
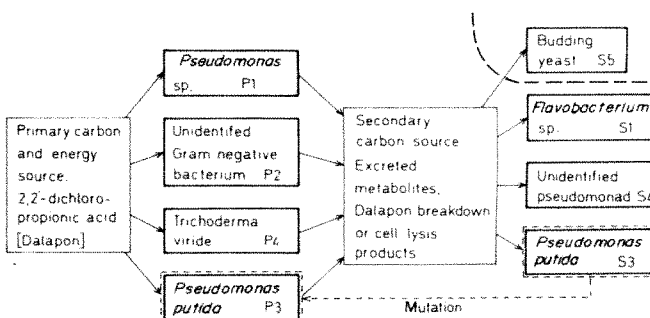
Fig. 1 The structure of the microbial community growing on 2,2'-dichloropropionic acid. The chemostat enrichment cultures had a volume of 0.5 l and were operated at a dilution rate (*D*) of $0.015\ h^{-1}$. The growth medium consisted of (g l^{-1}): K$_2$HPO$_4$, 1.5; KH$_2$PO$_4$, 0.5; (NH$_4$)$_2$SO$_4$, 0.5; MgSO$_4$.7H$_2$O, 0.2; Dalapon, mixed Na and Mg salts (Dow Chemical Company, UK) 1.65; pH 7.3. Several enrichments were set up using as inocula, *a*, culture effluent from a soil column containing a sandy loam (*p*H 4.8) which had been pretreated with Dalapon (the column had been perfused with Dalapon-containing minimal medium for several weeks before some of the culture effluent was used as the inoculum) and *b*, soil, either sandy loam or sandy chalk, which had or had not been pretreated with Dalapon, added directly to the chemostat. The enrichment cultures were monitored for biomass (culture absorbance and dry weight), species composition, pH and residual Dalapon concentration (determined by measuring chloride ions in solution by modification of Mohr's method[6]).

*Present address: Department of Applied Biology, University of Wales Institute of Science and Technology, King Edward VII Avenue, Cardiff CF1 3NU, UK.

The remaining six-membered association was exceedingly stable, however. One of the isolated communities was cultured for over 7,500 h at constant growth rate and environmental conditions with no qualitative change in the species composition. A second six-membered community has been continuously grown in a Dalapon-limited chemostat for over 13,500 h, which corresponds to approximately 3,700 generations. Throughout this period, and in spite of repeated environmental perturbations and changes in the specific growth rate, only one significant change in the community's character has been detected and this is described later. This particular community has been subjected to regular changes in dilution rate, such that seven continuous cycles of increasing and decreasing growth rate have been completed.

In the first cycle (Fig. 2) the dilution rate was increased by small increments and steady-state cultures were established at each value. At growth rates below $D=0.25$ h^{-1} the culture pH remained approximately constant at 3.3. The reduced pH in the poorly buffered growth medium and in the absence of fermenter pH control was due to the complete breakdown of the available Dalapon and the concomitant production of HCl. With a further slight increase in the dilution rate there was a marked rise in the culture pH to 6.1. This shift in pH was simultaneously accompanied by an increase in the unused Dalapon concentration and a decrease in the culture biomass. Further increases in the dilution rate, up to a maximum of 1.1 h^{-1}, caused further increases in the pH up to a maximum of 6.9. When the dilution rate was gradually decreased to complete the first cycle of dilution rate changes, a reverse hysteresis-like pattern of pH change was observed with the rapid phase of pH decline occurring at dilution rates between $D=0.19$ and 0.17 h^{-1}.

With successive cycles of changing dilution rate, the rate of pH change decreased until, by the seventh cycle, the pH change took place gradually over almost all the dilution rate range examined; moreover, in these later cycles the culture pH did not fall to the original low values. These observations indicated that to some extent the microbial community gradually stabilised, without qualitative changes in the species composition of the association, to dampen and minimise the environmental changes which occurred with changes in the dilution rate. The overall effect was to reduce the potentially deleterious influence of rapid pH fluxes.

During the increasing phase of the first dilution rate cycle (Fig. 2), the sudden pH shift was accompanied by an immediate development of heavy growth on the walls throughout the chemostat vessel. When the dilution rate was decreased the wall growth disappeared once the pH values had dropped to ~ 4.0. In later cycles the appearance of wall growth was much more gradual, probably because the change in pH was correspondingly less rapid. In all these latter cycles, however, heavy wall growth became reestablished when the pH had risen to greater than 5.0. The composition of the wall growth was complex, but it contained all the individual species of the microbial community originally present in free suspension. The attached surface growth was dominated by the mycelia of *Trichoderma viride*, however, and this contrasted markedly with the microflora in suspension at the same dilution rates, from which the fungus was virtually absent. At the lower dilution rates, and hence at low pH values, hyphae of *T. viride* were dispersed homogeneously in the culture. The development of wall growth prevented washout of the culture at dilution rates exceeding the maximum specific growth rate of the community and effectively preserved the integrity of the community in such circumstances. Submerged liquid batch and agar plate cultures of *T. viride* growing on Dalapon proved that the fungus readily metabolised the herbicide and grew vigorously at low pH values whereas at pH>6.0
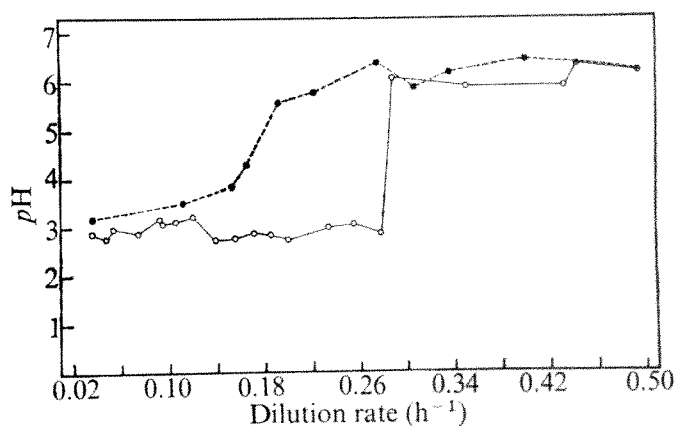
Fig. 2 The change in steady-state culture pH for the six-membered microbial community growing in a Dalapon-limited chemostat during the first cycle of increasing and decreasing the dilution rate. ○, Steady-state pH values in the increasing dilution rate phase; ●, pH values during the decreasing dilution rate phase. Heavy wall growth dominated by *T. viride* was established at growth rates in excess of $D = 0.28$ h^{-1}, during the increasing dilution rate phase, and had disappeared between $D = 0.19$ and 0.17 h^{-1} during the decreasing stage.

growth was extremely sparse. Thus fungal growth on the internal surface of the chemostat seemed to be a response to unfavourable conditions, namely high pH values, which developed at high dilution rates, and an attempt to create a favourable microenvironment within the chemostat where low pH conditions could be established and maintained.

New primary Dalapon utiliser

The significant change in the nature of the microbial community, first noticed after 2,900 h of continuous growth, was the appearance of a fourth primary Dalapon utiliser, subsequently identified as *P. putida* (P3) and identical in all other respects to the secondary utiliser, *P. putida* S3. It seemed possible that P3 had arisen as a mutation of S3 during the prolonged period of growth under Dalapon-limited conditions. The evolvant was retained within the community, together with its parent S3, and did not seem to unbalance the system. It is interesting that the selection of another primary utiliser should have occurred in spite of the presence of the three original primary utilisers and the inevitable increase in competition for the growth-limiting nutrient. Indeed the presence of three or four primary utilisers apparently contradicts the kinetic theory of enrichment cultures[5,6] which predicts that the organism with the greatest affinity for the limiting nutrient (that is, the lowest saturation constant) or the highest maximum specific growth rate will be the most successful in competing with others. In open systems, such as continuous-flow cultures, the most competitive organism ought to cause the elimination of the less competitive population. Our observations strongly indicate that several, so far unidentified, interactions occur between the individual populations of the whole community and these serve to stabilise free competition between the primary utilisers. The catabolism of Dalapon by hydrolysis to pyruvic acid and free chloride occurs as a result of the activity of the enzyme dehalogenase[3]. Thus *P. putida* P3 could have arisen either by the acquisition of Dalapon-metabolising capability by plasmid transfer from one of the original primary utilisers, or, by the selection of a mutant having a modified dehalogenase activity enabling it to metabolise Dalapon.

The latter hypothesis has been verified in monoculture chemostat studies. *P. putida* S3, isolated from the microbial community and purified could not grow on Dalapon in
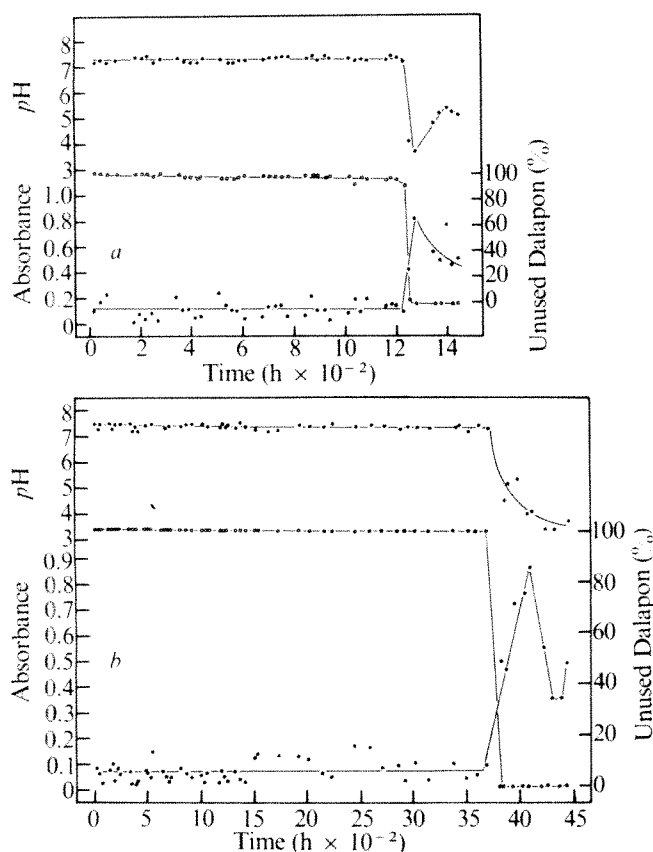
Fig. 3 *a* and *b*, Two experiments demonstrating the evolution of Dalapon-metabolising *P. putida* P3 from the non-Dalapon-metabolising *P. putida* S3. See text for details. Lower ●, culture absorbance (E_{625}); ○, unused Dalapon as a percentage of the concentration in the inflowing medium. Top ●, culture *p*H.

a range of environmental conditions. A culture of *P. putida* S3, grown on sodium propionate, was inoculated into a chemostat containing a minimal growth medium with the carbon source as the growth-limiting substrate and operating at a dilution rate of 0.05 h⁻¹. The carbon source was a mixture of sodium propionate (0.07 g carbon l⁻¹) and Dalapon (0.35 g carbon l⁻¹) which gave a carbon ratio of 1 : 5. S3 was initially able to grow only on the propionate carbon and thus could be retained in the open culture system and at the same time be continuously exposed to a non-utilisable carbon source. Accordingly evolvants with the capacity to utilise the additional carbon source might be expected to have a considerable selective advantage if they could acquire the capacity to metabolise Dalapon.

Figure 3*a* and *b* shows the results of two such evolution experiments. In the first experiment (Fig. 3*a*) a steady-state culture of strain S3, growing solely on propionate was established with a culture absorbance of approximately 0.1, a situation that was maintained for just over 1,200 h of continuous growth in constant environmental conditions. During the first stage of the experiment the culture *p*H remained constant at its initial value of 7.4, suggesting that no Dalapon was being metabolised. This conclusion was confirmed by measuring free chloride in the culture and calculating the percentage of Dalapon which had been metabolised. After 250 h of growth a small number of Dalapon-metabolising colonies were detected on dilution plates. Initially the detection of such colonies was variable but with time their frequency of detection increased until they were found in every sample taken from the chemostat. The presence of Dalapon-metabolising individuals within the population did not immediately result in their competitive selection and displacement of the S3 population.

After 1,200 h of growth, however, such an event did occur and was manifest as an approximately fivefold increase in the culture absorbance, a result which would be anticipated if the evolvant population were able to utilise the additional carbon source. Metabolism of Dalapon was indicated by a concomitant decrease in culture *p*H and an increase in free chloride in the culture. Within 30 h the evolvant **P3** succeeded in establishing a new steady-state culture and completely displacing the original S3 strain. In the new steady state the Dalapon concentration in the chemostat was reduced to zero. A similar sequence of events has been observed in five other experiments, the only variation being the time at which the Dalapon-metabolising population managed to replace the parent strain. In a second experiment (Fig. 3*b*) for example, the mutational event leading to the selection of the P3 strain occurred much later, after 3,750 h of continuous growth, in conditions identical to those of the first experiment.

Preliminary dehalogenase analyses have shown that during the first 3,750 h of growth of *P. putida* S3 (Fig. 3*b*) there was no dehalogenase activity with a specificity towards Dalapon. After the mutational event and the establishment of the new steady-state culture however, Dalapon-specific dehalogenase activity was detected (E. Senior, C. Larvin, J. H. Slater and A. T. Bull, unpublished). In addition, we have obtained evidence which shows that the dehalogenase originally possessed by *P. putida* S3 had a specificity towards 2-monochloropropionic acid and that the evolvant P3 has a markedly increased specific activity of the 2-monochloropropionic acid-specific dehalogenase (J. H. Slater, D. E. Sanders and A. T. Bull, unpublished). Growth experiments have indicated that *P. putida* S3 will grow slowly on 2-monochloropropionic acid as its sole carbon source, whereas P3 grows vigorously on this substrate.

Enzyme evolution

Experimental enzyme evolution is an area that has excited growing interest over the past decade and elegant demonstrations of the evolution of enzymes with new specificities and activities have been provided by studies of the aliphatic amidases of *P. aeruginosa*[7] and the pentitol dehydrogenases of *Klebsiella aerogenes*[8]. The chemostat offers considerable scope for investigating the evolutionary potentialities of microorganisms, principally because it enables the imposition of precise selection pressures on a population. The putative evolution of *P. putida* dehalogenase is a further vindication of this experimental approach and it is worth emphasising that it was achieved without the intervention of specific mutagens. Moreover, we would like to emphasise the value of mixed microbial populations in enzyme evolution studies in two important respects. First, within such a mixed system multiple selective pressures can operate which are analogous to those occurring in real ecosystems. Second, in the microbial community we have begun to analyse it is clear that the biochemical and physical nature of the whole system provides a permissive environment for metabolic evolution. In particular, the secondary species, by virtue of the fact that they are continuously supplied with assimilable substrates by the activity of the primary utilisers, are retained within the community and thus are continuously exposed to a non-metabolisable compound. Indeed it is clear that for enzyme evolution experiments in monoculture a metabolisable 'carrying' substrate has to be provided to ensure that the single species population continues to grow and be exposed to the non-metabolisable substrate for long periods. We are now attempting to elucidate the precise nature of the dehalogenase in the parent *P. putida* and its relationship to the evolvant strains and to explore the extent to which the catalytic activity of this enzyme can be modified to deal with different chlorinated aliphatic

acids. E.S. thanks the SRC for a research studentship. We thank the Dow Chemical Company for supplying Dalapon; A. Gallenkamp and Co. Ltd, for assistance with fermentation systems and Professor H. Veldkamp and Dr W. Harder for discussions.

1 Veldkamp, H., and Jannasch, H. W., *J. appl. Chem. Biotechnol.*, 22, 105–123 (1972).
2 Veldkamp, H., *Sixth International Symposium on Continuous Culture* (edit. by Dean, A. C. R., Ellwood, D. C., Evans, C. G. T., and Melling, J.) (in the press, 1976).
3 Kearney, P. C., Kaufman, D. D., and Beall, M. L., *Biochem. biophys. Res. Commun.*, 14, 29–33 (1964).
4 Shewan, J., Hobbs, G., and Hodgkiss, W., *J. appl. Bact.*, 23, 379–390 (1960).
5 Powell, E. O., *J. gen. Microbiol.*, 18, 259–268 (1958).
6 Slater, J. H., Bull, A. T., and Senior, E., *Proc. Soc. gen. Microbiol.*, 3, 113 (1976).
7 Clarke, P. H., in *Evolution in the Microbial World* (edit. by Carlile, M. J., and Skehel, J. J.), (Cambridge University Press, London, 1974).
8 Hartley, B. S., in *Evolution in the Microbial World* (edit. by Carlile, M. J., and Skehel, J. J.), (Cambridge University Press, London, 1974).
9 Vogel, A. I., *A Textbook of Quantitative Inorganic Analysis*, third ed. (Longman, London, 1962).

Partial purification of rabbit aorta contracting substance-releasing factor and inhibition of its activity by anti-inflammatory steroids

F. P. Nijkamp, R. J. Flower*, S. Moncada & J. R. Vane

Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, UK

Rabbit aorta contracting substance-releasing factor (RCS-RF) is found in perfusates from guinea pig lungs during anaphylaxis. It has been identified as a small peptide which releases arachidonic acid from lung tissue, thus generating prostaglandin endoperoxides and thromboxanes and causing bronchoconstriction. Anti-inflammatory steroids block the release of arachidonate by RCS-RF. In this activity their relative potency is very similar to their relative anti-inflammatory potency, suggesting that the two actions may be related.

PIPER and Vane[1] detected the release of additional substances into the perfusate during anaphylaxis in isolated lungs from sensitised guinea pigs. Because of its activity they called one "rabbit aorta contracting substance" or "RCS". RCS was unstable at room temperature and its release was blocked by aspirin-like drugs. RCS has now been identified as a mixture of the prostaglandin cyclo-endoperoxides (PGG_2 and PGH_2) and thromboxane A_2 (TXA_2)[2,3]. The RCS activity is mainly due to TXA_2. Piper and Vane also found that after the RCS activity of the perfusate declined, another much more stable substance could be detected. This was called "RCS-releasing factor" or RCS-RF because effluent containing RCS-RF injected into the pulmonary artery of perfused lungs from unsensitised guinea pigs induced a release of "RCS". We have now isolated and partially purified RCS-RF, which seems to be a small peptide. RCS-RF stimulates the generation of PG endoperoxides and TXA_2 probably by releasing the precursor arachidonic acid from lung tissue. Steroid anti-inflammatory drugs block this releasing action and thus prevent the generation of TXA_2, whereas non-steroid anti-inflammatory drugs block the generation of TXA_2 by a direct effect on the prostaglandin cyclo-oxygenase enzyme. The rank order of nine steroid drugs as inhibitors of RCS-RF activity correlates well with their rank order as anti-inflammatory substances.

Isolation, partial purification and assay

The assay procedure for RCS-RF and SRS-A like activity is described in the legend to Fig. 1. Lungs were removed from male guinea pigs (200–300 g) which had been sensitised

---

*To whom all correspondence should be addressed.

to ovalbumin (100 mg subcutaneously and 100 mg intra-peritoneally) 2–4 weeks earlier. The lungs were perfused through the pulmonary artery with Krebs' solution at 10 ml min^{-1} as previously described[1]. Anaphylactic shock was induced by injection into the pulmonary artery of 10 mg ovalbumin and the perfusate was collected for 10 min (about 100 ml). Effluent from 5–10 sensitised lungs was pooled, centrifuged (5,000g for 30 min) to remove cells, filtered and stirred for 30 min with 30 g l^{-1} Amberlite XAD-2 (BDH) to adsorb the RCS-RF activity. After filtering under vacuum to remove the Krebs' solution the Amberlite XAD-2 was washed with small volumes of distilled water,

Fig. 1 For bioassay of RCS-RF the effluent from unsensitised lungs was allowed to superfuse a rabbit aortic strip, a rat fundic strip, a rat colon and guinea pig trachea in cascade[1]. Sensitivity and selectivity of the tissues were increased by an infusion into the lung effluent of "combined antagonists"[1] to prevent the actions of histamine, acetylcholine, catecholamines and 5-hydroxy-tryptamine as well as indomethacin (1 μg ml^{-1}) to prevent endogenous synthesis of prostaglandins by the tissues[4]. SRS-A-like activity of samples was estimated using the isolated guinea pig ileum[5] preparation, hyoscine hydrobromide (10^{-7} g ml^{-1}) and, where appropriate, mepyramine maleate (10^{-7} g ml^{-1}) was added to the organ bath to prevent spontaneous activity and render the preparation insensitive to histamine[5]. PP, Pulmonary artery perfusion pressure; RA, rabbit aortic strip; RSS, rat (fundic) stomach strip; RC, rat colon; GPT, guinea pig trachea; OT, injection directly over the tissues; TL, injection into the pulmonary artery; AA, arachidonic acid.

Table 1 Partial purification of RCS-RF

Extraction procedure	Weight of sample	RCS-RF-like activity (Total units)	Yield (%)	Specific activity (U mg^{-1})	Purification factor
None. Original effluent lyophilised*	6.64 g	3,534.23	100	0.53	0
Amberlite XAD-2 extraction	245.6 mg	1,473.77	41.70	6.06	11.43
Diethyl ether extraction	236.9 mg	1,208.71	34.20	5.10	9.62
Paper chromatography	0.22 mg	692.12	19.58	3,146.00	5,935.84

The extraction efficiencies vary from experiment to experiment and the figures quoted in this table should only be regarded as a guide.
*Starting material was 570 ml effluent.

and the RCS-RF recovered from the resin with 40-50% yields (see Table 1) by washing with 1.5 l ethanol–water (80 : 20, v/v). The ethanol–water extract was evaporated *in vacuo* in a rotary film evaporator, and the dried residue was washed three times with dry diethyl ether to remove lipids (thin-layer chromatography of the ether washings revealed substantial quantities of fatty acids, prostaglandins and some phosphatides). After ether washing the residue was dissolved in a small aliquot of distilled water: this preparation, referred to as the "crude fraction", possesses potent RCS-RF activity and is stable for at least 6 weeks when kept frozen.

For further purification, the crude fraction was streaked on to a silica gel-impregnated paper chromatogram (Whatman SG81) which was developed by ascending chromatography in ammonia–*n*-propanol–water (30 : 60 : 10, v/v). After developing to a distance of 20 cm, the chromatogram was removed, dried and cut into 1-cm horizontal sections. There were eluted with small volumes of distilled water and the biological activity assayed.

About 80% of the RCS-RF activity was associated with a single band of R_f 0.7 (see Fig. 2). A small amount of RCS-RF activity was also recovered from a zone running ahead of the major peak and sometimes small amounts were detected close to the origin (see Fig. 2). In addition to the RCS-RF activity, there were several zones containing activity which contracted the guinea pig ileum, indicating SRS-A like activity. One of these corresponded to the bulk of the RCS-RF activity, and others to the minor RCS-RF zones. We regarded the major zone of releasing activity as "RCS-RF" and referred to this fraction as the "partially purified" fraction. The biological activity of the partially purified fraction is shown in Fig. 1. Partially purified RCS-RF has no "direct" effect over the smooth muscle organs but when injected into isolated perfused lungs from guinea pigs it provoked an immediate increase in perfusion pressure, a bronchoconstriction (not shown), and a release of RCS; prostaglandins E_2 or $F_{2\alpha}$ were seldom detected.

The partially purified fraction was stable for several weeks when kept frozen. The releasing activity of RCS-RF was potentiated three- to fivefold by concomitant injections of 10 μg of an inhibitor of carboxydipeptidase activity BPP-5A (ref. 6). We defined "1 unit" (U) of RCS-RF as that amount which when injected into the lung released the same amount of RCS-activity (TXA$_2$) as 1 μg of arachidonic acid.

Physicochemical properties

RCS-RF is insoluble in ether, chloroform, ethyl acetate, acetone and ethanol but is more soluble in methanol (about 70% of its water solubility). It is destroyed by drastic *p*H changes, being more susceptible to acid rather than alkaline *p*H values. For example, a 1-h incubation with 1 N HCl completely inactivated RCS-RF but a 1-h incubation with 10 N NaOH still left about 20% of the activity. RCS-RF loses 75% activity after boiling for 10 min, but was unaffected by boiling for 1 min. Passage through a dialysis

membrane suggests a small molecule of <5,000 molecular weight.

The activity of RCS-RF is not reduced by incubation with phospholipase A, C or D (Sigma) (100 μg enzyme per ml) but was reduced slightly (~15%) by incubation with trypsin (Sigma) (500 μg enzyme per ml). In two out of four experiments, incubation with arylsulphatase (Sigma) (500 μg ml^{-1}) produced 50–60% inactivation; SRS-A is also inactivated by this enzyme. Incubation with carboxypeptidase B (Sigma) (500 μg enzyme per ml for 1 h at 37 °C) or aminopeptidase M (Sigma) (100 μg enzyme per ml for 1 h at 37 °C) resulted in an 80–85%, or complete inactivation (respectively), strongly suggesting that RCS-RF is a peptide. When a freshly developed paper chromatogram was sprayed with ninhydrin reagent, which reacts with α-amino acyl groups of amino acids, the zone containing RCS-RF was only one of several ninhydrin-positive areas. RCS-RF was not found in the perfusate of non-sensitised lungs challenged with antigen and, unlike RCS, was not released by mechanical trauma. The limited physicochemical data obtained so far are compatible with RCS-RF being a peptide of less than 10 amino acids.

Effects of inhibitors

The release from sensitised lungs of RCS-RF by antigen was not prevented by aspirin (200 μg ml^{-1}), in the perfusing fluid, indomethacin (2 μg ml^{-1}), dexamethasone (2 μg ml^{-1}), disodium cromoglycate (20 μg ml^{-1}), colchicine (5 μg ml^{-1}), diethyl carbamazine (1 mg ml^{-1}), or mepacrine (20 μg ml^{-1}). Aspirin, indomethacin, mepacrine and dexamethasone (same concentrations) all blocked the release of TXA$_2$ from the

Fig. 2 Biologically active material recovered from the paper chromatogram. Top panel RCS-RF-like activity assayed as shown in Fig. 1; bottom panel SRS-A-like activity, assayed as described in Fig. 1.
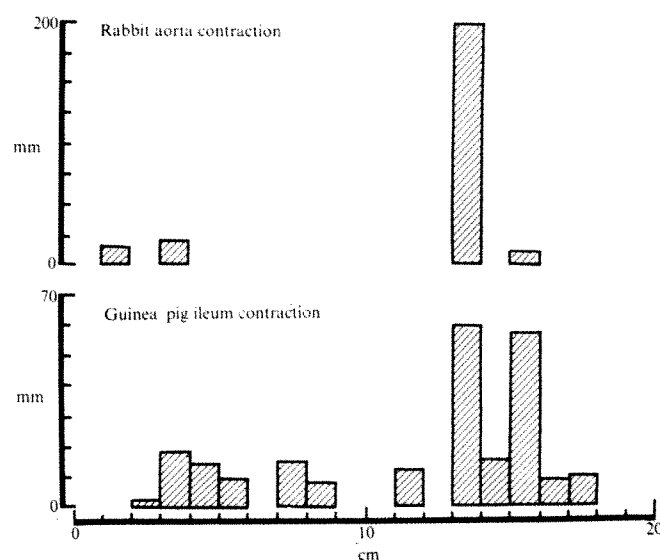
		Table 2	ID_{50} Values for some anti-inflammatory steroids	
Steroid	ID_{50} (nmol min^{-1})	Inhibition (%) of arachidonic acid at ID_{50}	Relative (molar) potency	Relative anti-inflammatory potency*
Dexamethasone	34.4	− 0.3	35.1	25.0
Betamethasone	36.5	− 7.3	33.1	25.0
Triamcinolone	106.1	+31.1	11.4	∼5.0
Fludrocortisone	133.0	−21.9	9.1	10.0
Prednisolone	177.1	− 0.1	6.8	4.0
Prednisone	185.8	− 2.3	6.5	4.0
Corticosterone	968.9	−32.84	1.3	0.35
Cortisone	1,206.8	+ 9.12	1.0	0.8
Hydrocortisone (cortisol)	1,207.0	− 11.9	1.0	1.0

The following drugs were used: dexamethasone sodium phosphate (Decadron, Merck Sharp and Dohme) or dexamethasone (Sigma); betamethasone (Betsolan, Glaxo); triamcinolone acetonide (Kenalog, Squibb); fludrocortisone (Sigma); prednisolone acetate (Deltastab, Boots); prednisone (Koch-Light); corticosterone (Sigma); cortisone (Sigma); and hydrocortisone sodium succinate (Efcortelan, Glaxo). All steroids were infused for 15 min, and the amounts indicated are calculated as the base. Apart from the soluble esters, steroids were dissolved in ethanol and diluted 9:1 (v/v) with 0.9% saline. Each result is the mean of three or four observations.
*Data taken from ref. 8.

lung provoked by RCS-RF. All these substances except dexamethasone also prevented the conversion of exogenous arachidonic acid into TXA₂, confirming that they act by a direct block of cyclo-oxygenase[4]. Dexamethasone, however, blocked the release of TXA₂ by RCS-RF without affecting the conversion of exogenous arachidonic acid to TXA₂, implying a different site of action. We therefore examined the effects of other steroids in this system. Table 2 shows that all the steroids tested blocked the releasing activity of RCS-RF. Furthermore, their relative activity in this test closely paralleled their reported anti-inflammatory activity. The steroids had variable but generally only small

Mechanism of action and site of steroid blockade

Infusion of arachidonic acid into perfused guinea pig lungs leads to an immediate synthesis of TXA₂ suggesting that substrate availability is a rate-limiting step in this system. RCS-RF may, therefore, act by liberating arachidonate from some intracellular pool, and steroids could be preventing this effect. To test this possibility, we perfused isolated lungs with Krebs' solution containing 5,8,11,14-eicosatetraynoic acid (TYA)—an acetylenic analogue of arachidonate—at 10 μg ml^{-1}. This prevented metabolism of any

					Basal	RCS-RF alone	RCS-RF + steroid	RCS-RF alone	Net release by RCS-RF
Experiment	Steroid	Dose (nmol min^{-1})	Sample collection period (min)	RCS-RF injected	Arachidonic acid efflux from lung (μg)
1	None	—	7	5 U (Crude)	0.15	0.22	—	—	0.07
2	None	—	7	5 U (Partially purified)	0.18	0.24	—	—	0.06
3	None	—	7	5 U (Partially purified)	0.26	0.31	—	—	0.05
4	Betamethasone	178.3	5	8 U (Crude)	1.13	5.99	0.78	—	4.86
5	Dexamethasone	178.3	7	8 U (Crude)	0.38	1.63	0.43	1.24	1.35
6	Fludrocortisone	727.7	7	4.5 U (Purified)	0.78	1.16	0.65	0.78	0.38
7	Hydrocortisone	5,517.2	7	4.5 U (Partially purified)	—	1.56	0.52	1.92	∼1.04

Table 3 Release of arachidonate from lungs by RCS-RF

Lung effluent was collected for 5–7 min following an injection of RCS-RF either in the presence or absence of blocking doses of steroids. The pH of the effluent was adjusted to 3.0 with 1 N HCl and the arachidonate extracted with ether, a small amount of 1-¹⁴C-arachidonate was added as a tracer. After concentration of the sample *in vacuo* the fatty acids were methylated with diazomethane and separated on thin-layer chromatography silica gel plates impregnated with 10% AgNO₃ (Uniplates, Anachem) by development in n-hexane, diethyl ether, acetic acid (50:50:1 v/v). The methyl arachidonate zone was located with a radiochromatogram scanner and eluted with ether, and the arachidonate content was estimated with a Pye Unicam series 104 gas chromatograph, using a 5-foot column packed with 10% DEGS on Gas-Chrom Q. With an oven temperature of 195 °C methyl arachidonate was easily detected by the flame ionisation detector, having a retention time of about 8 min. The labelled arachidonic acid was used to correct for losses during sample manipulation. Large doses (5–10 U) of RCS-RF were used to ensure that easily detectable amounts could be liberated, and correspondingly higher doses of steroids were used to compensate for this.

effects on the conversion of arachidonic acid to TXA₂, which were unrelated to their anti-RCS-RF activity. The release of TXA₂ by arachidonate, for example, was unaffected by dexamethasone, potentiated by triamcinolone and depressed by fludrocortisine. These effects could be related to the vehicle in which the steroid was dissolved; for ethanol (used to dissolve some of the steroids) in concentrations greater than 10 μl ml^{-1} depressed the release of TXA₂ by arachidonate.

There were two other notable features concerning the block of TXA₂ release by anti-inflammatory steroids; first, a maximal effect was obtained only after infusing each steroid for 10–20 min. Second, the inhibition produced by steroids was easily overcome by increasing the amount of RCS-RF. Thus, the ID_{50} of a steroid will depend on the concentration of the agonist.

released arachidonate by the lung, TYA being an inhibitor of both the cyclo-oxygenase and lipoxygenase pathways of arachidonate metabolism[9,10]. Table 3 shows that both "partially" purified and "crude" RCS-RF liberated arachidonate from lungs and that this release was blocked when any of the steroids was present. Thirty minutes after stopping the steroid infusion, arachidonate release by RCS-RF returned, indicating that the effects of the steroids in this system were relatively short lived and easily reversed.

The actual amounts of arachidonate released from lungs varied greatly between experiments (see Table 3, experiments 4–7) but were relatively consistent in single lungs (experiments 1–3). Although the samples were collected for 5–7 min it should be realised that the evoked release of arachidonate usually takes place within 1 min giving an effect similar to a bolus injection of the substrate. It is

interesting to note that there is a continuous basal release of arachidonate from the isolated lungs. This was generally found to increase with time (experiments 1–3, Table 3). In one experiment (No. 4) at least, betamethasone actually lowered the stimulated release below basal levels. This observation correlates with an effect frequently observed during steroid infusion—a decline in tone of the bioassay organs, suggesting a decreased basal TXA_2 output.

What is the source of the arachidonic acid liberated by RCS-RF? Although amounts of free arachidonic acid in many tissues are small (see ref. 11) considerable quantities are present in ester form in the neutral lipid or phosphatide fractions. Several observations (see ref. 11) indicate that phospholipids are the chief source of arachidonate in cells, so that the phospholipase enzyme system may be stimulated in some way by RCS-RF. Some evidence for this possibility has already been obtained in this laboratory (G. J. Blackwell, R.J.F. and F.P.N., unpublished). Specifically labelled (2'-(1-^{14}C)-oleoyl) lecithin is partially (~ 1–3%) hydrolysed when perfused through the isolated lung—presumably by a phospholipase A_2. The rate of hydrolysis is doubled when concomitant injections of 5 U RCS-RF were given, and decreased when infusions of dexamethasone or betamethasone (178.8 nmol min^{-1}) were made. Phospholipase A_2 activity in guinea pig lung homogenates, however, was not affected by the anti-inflammatory steroids, suggesting that these drugs only have a blocking action in intact cells.

RCS-RF and anaphylaxis

We do not know whether RCS-RF formation and release contributes to anaphylaxis in guinea pigs. It does not seem to have any intrinsic smooth muscle activity in the doses we used here. Our finding that it cannot be recovered from the lungs of unsensitised animals or from unchallenged sensitised lungs would suggest that it is not stored in the lung, but is elaborated in response to the immunological stimulus. Thus, if it does have a role in anaphylaxis it could be among those substances classified as "primary mediators" by Austen and Orange[12]. If it is a peptide, RCS-RF cannot be the SRS-A of Orange *et al.*[7], a conclusion reached earlier by Piper and Vane[1] on the basis of biological evidence. Although there seem to be similarities between SRS-A and RCS-RF, any formal relationship between the two compounds (as suggested by the arylsulphatase inactivation experiment) awaits further clarification.

Of great interest is the finding that anti-inflammatory steroids block the RCS-RF induced release of thromboxane. Even though they do not inhibit the cyclo-oxygenase[13] there has been increasing evidence that one steroid (hydrocortisone) interferes in some way with prostaglandin biosynthesis in several other systems[14–17]. The mechanism proposed by Lewis and Piper[14] to account for hydrocortisone action, that of an inhibition of a cellular transport mechanism, does not operate here since arachidonate consistently reversed the steroid block. Instead, our results point to a mechanism of action much earlier in the biosynthetic pathway—the inhibition of substrate liberation from intracellular stores. A similar hypothesis has already been advanced by Gryglewski and his colleagues[15]. This store is probably the phosphatide fraction of the cell, and the enzyme responsible for liberation of the acid, the cell membrane phospholipase A_2. Bradykinin is another peptide which releases RCS from the lung[18] and, like RCS-RF, its activity is potentiated by BPP-5A. We find, however, that only RCS-RF activity is steroid sensitive, implying that it activates a very selective enzymatic step.

The order of potency of the steroids in inhibiting RCS-RF activity is striking similar to the order of anti-inflammatory potency, and the possibility that inhibition of substrate release is a general method by which anti-inflammatory steroids can interfere with the synthesis of prostaglandin (or other arachidonate metabolites) must be considered.

1 Piper, P. J., and Vane, J. R., *Nature*, 223, 29–35 (1969).
2 Svensson, J., Hamberg, M., and Samuelsson, B., *Acta. physiol. scand.*, 94, 222–228 (1975).
3 Bunting, S., Moncada, S., and Vane, J. R., *Br. J. Pharmac.* 57, 462P–463P (1976).
4 Vane, J. R., *Nature new Biol.*, 231, 232–235 (1971).
5 Brocklehurst, W. E., *J. Physiol., Lond.*, 157, 416–435 (1960).
6 Ferreira, S. H., Bartelt, D. C., and Greene, L. J., *Biochemistry*, 9, 2583–2594 (1970).
7 Orange, R. P., Murphy, R. C., and Austin, K. F., *J. Immun.*, 113, 316–322 (1974).
8 Sayers, G., and Travis, R. H., in *The Pharmacological Basis of Therapeutics*, fourth ed. (edit. by Goodman, L. S., and Gilman, A.), 1604–1642 (Macmillan, New York, 1970).
9 Hamberg, M., and Samuelson, B., *Biochem. biophys. Res. Commun.*, 61, 942–949 (1974).
10 Hamberg, M., and Samuelsson, B., *Proc. natn. Acad. Sci. U.S.A.*, 71, 3400–3404 (1974).
11 Flower, R. J., and Blackwell, G. J., *Biochem. Pharmac.*, 25, 285–291 (1976).
12 Austen, K. F., and Orange, R. P., *Am. Rev. respir. Dis.*, 112, 423–436 (1975).
13 Flower, R. J., *Pharmac. Rev.*, 26, 33–67 (1974).
14 Lewis, G. P., and Piper, P. J., *Nature*, 254, 308–311 (1975).
15 Gryglewski, R. J., Panczenko, B., Korbut, R., Grodzinska, L., and Ocetkiewicz, A., *Prostaglandins*, 10, 343–355 (1975).
16 Tashjian, A. H., Voelkel, E. F., McDonough, J., and Levine, L., *Nature*, 258, 739–741 (1976).
17 Kantrowitz, F., Robinson, D. W., McGuire, M. B., and Levine, L., *Nature*, 258, 737–739 (1976).
18 Vargaftig, B. B., and Dao Hai, N., *J. Pharm. Pharmac.*, 24, 159–161 (1975).

letters to nature

Cosmic rays near the galactic centre

THE galactic centre (GC) is characterised by objects of high activity: there are powerful sources of infrared radiation, radio waves and X rays, and evidence for a small region (~ 1 pc) of extremely high luminosity (~ $10^8 L_\odot$). The presence of a ring of gas (very largely molecular hydrogen)[1] at a distance of ~ 300 pc from the GC also indicates violent activity. It is reasonable to assume that there is also an increased cosmic-ray intensity at least for energies ≲ 10 GeV. The presence of protons and electrons having such energies together with high gas densities and high starlight intensities provides the ingredients for significant γ-ray fluxes. Here we examine the implications of recent measurements of the γ-ray flux from the region of the GC, with particular reference to the 300-pc ring.

The most comprehensive γ-ray data so far obtained are those from the SAS-II satellite (see ref. 2). The earlier data did not give strong evidence for a peak at l^{11} ~ 0° but a very recent reanalysis by Fichtel *et al.*[3] shows a rather impressive peak (see Fig. 1) which offers the possibility of deriving the cosmic-ray intensity in the GC region. It should be remarked, however, that the only other precise cosmic γ-ray experiment, that on the COS B satellite[4], does not show a peak near l^{11} = 0°. The analysis so far is described as "preliminary", however, and we are of the view that the peak is still very probably genuine.

Although the SAS-II peak looks rather broad (FWHM ≃ 6°) most of this is instrumental and it seems that the bulk of the radiation is coming from within about ±2° of the GC. The distribution thus seems to be both narrower and of higher
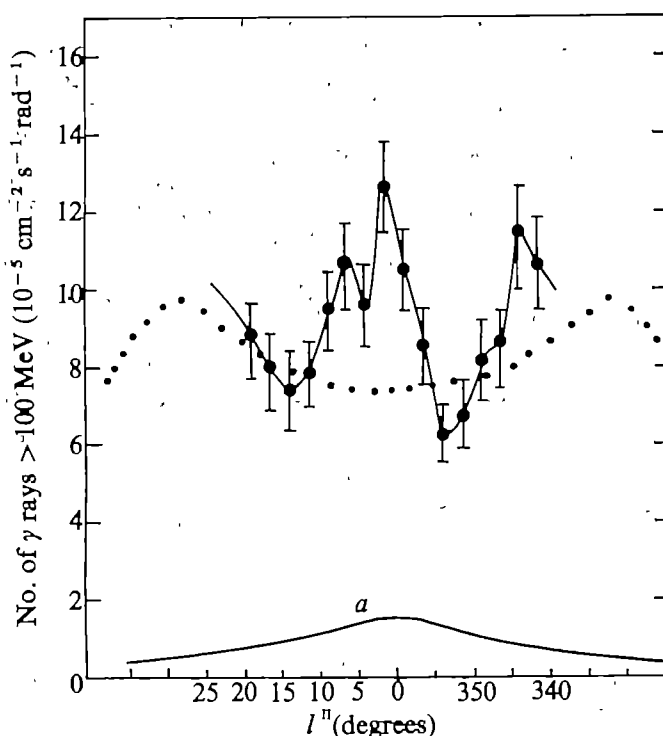
Fig. 1 The flux of γ rays above 100 MeV for longitudes close to the GC, from Fichtel *et al.*[3]. ● ● ●, Estimate of the contribution for $R > 350$ pc, where we assume that the cosmic-ray intensity is proportional to the mean surface gas density. *a*, Contribution from electrons by way of the inverse Compton effect, assuming a constant electron flux throughout the Galaxy[4].

intensity than estimated for inverse Compton interactions by Dodds *et al.*[5] (see Fig. 1) for the case of an electron intensity near the GC similar to that locally (and with a full width for the electron distribution of 500 pc). There is thus evidence for an increase in the cosmic-ray intensity near the GC.

The original SAS-II data have been analysed by many authors and it has been shown that there is almost certainly a gradient of 1–10-GeV cosmic-ray protons in the Galaxy. Our view is that the firmest evidence comes from the analysis of measurements towards the galactic anti-centre[6] but the most reasonable interpretation of γ rays from the inner Galaxy is that there is a gradient there too, with the cosmic-ray intensity roughly following the average surface gas density[7] and the surface density of supernova remnants and pulsars[8,9] (at least for $R \gtrsim 4$ kpc). Such a correlation would be expected if the cosmic-ray lifetime is roughly constant for $R \gtrsim 4$ kpc, in which case the intensity will be proportional to the generation rate.

Near the GC we might expect the generation rate to be again roughly proportional to the gas density, and in what follows we examine whether this is the case.

The SAS-II data of Fig. 1 indicate a flux of γ rays from within a few degrees of the GC of $\sim 6.7 \times 10^{-4}$ cm^{-2} s^{-1}. This value has been derived using as a datum for non-centre flux the distribution shown dotted in Fig. 1; in the calculation the contribution from the rest of the Galaxy followed from the assumption that the cosmic-ray intensity is proportional to mean surface gas density.

Locally, the escape times for electrons and protons are roughly equal and therefore the e–p ratio is expected to be the same as at generation. In the GC region the lifetime against energy loss is about the same for electrons and protons (see later) and so we expect their intensities to be enhanced by the same factor (F) over their local values. The procedure is to calculate F so that the measured flux of γ rays at the Earth is reproduced.

The first contribution arises from the decay of π^0s produced by protons and heavier cosmic-ray nuclei interacting with the

interstellar matter at the GC. For the mass of gas and its distribution we use the measurements of Scoville and Solomon[10] which indicate that there is about $5 \times 10^7 M_\odot$ of molecular hydrogen distributed within a radius of ~ 350 pc and a volume of $\approx 10^7$ pc^3. Assuming that the cosmic-ray spectrum has the same spectral shape as that locally, the expected flux of γ rays above 100 MeV at the Earth is $\sim 6.8 \times 10^{-7}$ F cm^{-2} s^{-1}. Although there will be a contribution from inverse Compton interactions of electrons in the starlight field, in the present model this will be outweighed by the yield from bremsstrahlung interactions. We assume that the lifetime for escape of the cosmic-ray nuclei from the ring is significantly greater than 10^5 yr, the lifetime against interaction. The protons will then be in equilibrium with the secondary electrons from $\pi \to \mu \to e$ The electrons, too, are assumed to be absorbed before escaping from the ring, bremsstrahlung predominating (the mean lifetime is $\sim 2 \times 10^5$ yr at 1 GeV). Ramaty[11] has given the generation spectrum for electron secondaries and from this the secondary electron intensity has been derived. As an example, the derived intensity at 1 GeV is $\sim 12\%$ of that of primary protons in the ring. To this must be added $\sim 3\%$ for primary electrons giving a total electron intensity at 1 GeV in the ring of 15% of the proton intensity.

At electron energies of ~ 100 MeV ionisation losses predominate over bremsstrahlung and the γ rays at 100 MeV are mainly from π^0 decay. This implies $F \simeq 10$.

A value for F (to be denoted F_a) can also be calculated from non-thermal radio observations of the GC if we assume that there is a coupling between the cosmic-ray energy density and that in the magnetic field. Measurements at 408 MHz (ref. 12) suggest a brightness temperature of 100 K from the 300-pc molecular ring ($\sim 4 \times 10^{-4}$ sr) region, of which we estimate $\sim 25\%$ to be thermal[13]. Using a lifetime against bremsstrahlung of 2×10^5 yr, the equilibrium electron intensity has been calculated. Equating the calculated synchrotron flux to observations gives $F_a \simeq 12$. This is rather larger than the value of F calculated from the γ-ray flux but it is not inconsistent considering uncertainties in the electron lifetime and the likely underestimate of the percentage of thermal radiation.

The value $F = 10$ derived from the γ-ray analysis can be used to derive the relative generation rates in the ring, where the lifetime is $\sim 10^5$ yr, and locally, where we assume that the lifetime by escape is $\sim 2.5 \times 10^6$ yr. The ratio ring-to-local value is ~ 250; this can be compared with the ratio of gas densities, $\sim 200{:}1$. The nearness of the results adds some weight to our original contention.

We thank Drs T. M. Bania, J. L. Osborne, N. V. Scoville and A. W. Strong, Professors W. Butler Burton and P. M. Solomon, and Mr D. Dodds for discussions. Dr C. E. Fichtel and his colleagues kindly made available their new γ-ray data. D. M. W. thanks the SRC for a studentship.

A. W. Wolfendale
Diana M. Worrall

Physics Department,
University of Durham,
South Road,
Durham DH1 3LE, UK

1 Scoville, N. A., Solomon, P. M., and Jefferts, K. B., *Astrophys. J. Lett*, **187**, L63 (1974).
2 Fichtel, C. E., *et al.*, *Astrophys J*, **198**, 163 (1975).
3 Fichtel, C. E., *et al.*, *Proc. γ-ray Symp.*, *Goddard Space Flight Center, Maryland* (June 1976).
4 Bennett, K., *et al.*, *ibid.*
5 Dodds, D., Strong, A. W., Wolfendale, A. W., and Wdowczyk, J, *J. Phys. A.*, **8**, 624 (1975).
6 Dodds, D., Strong, A. W , and Wolfendale, A. W., *Mon. Not. R. astr. Soc.*, **171**, 569 (1975).
7 Bignami, G. F., and Fichtel, C. E., *Astrophys. J. Lett.*, **189**, L65 (1974).
8 Dodds, D., and Wolfendale, A. W., *Proc. γ-ray Symp. Goddard Space Flight.*
9 Stecker, F. W., *Nature*, **260**, 412 (1976).
10 Scoville, N. Z., and Solomon, P. M., *Astrophys. J. Lett*, 199, L105, (1976).
11 Ramaty, R., in *High Energy Particles and Quanta in Astrophysics* (edit by McDonald, F. B., and Fichtel, C. E.), 122 (MIT Press, Cambridge, 1974).
12 Green, A. J., *Astr. Astrophys. Suppl.*, **18**, 267 (1974).
13 Hirabayashi, H., *Publ. astr. Soc Jap.*, **26**, 263 (1974).

The 35-d X-ray profile of Her X-1

FOLLOWING the report from Ariel V (ref. 1) of substantial X-ray emission from Her X-1 midway through the extended low-state portion of its 35-d cycle, Jones and Forman[2] have reported a similar feature during another cycle, measured from Uhuru. They further suggest that this feature should occur regularly, consistent with models for the mass accretion disk which are inclined with respect to the binary orbital plane[3,4]. We report here the results of the analysis of > 500 d of Ariel V all-sky monitor (ASM) data, which support the view that the feature occurs regularly in the 35-d cycle.

The Ariel V ASM is barely sensitive to Her X-1 under the most favourable of viewing constraints (see ref. 5 for a complete description of the apparatus). Its sensitivity for accumulation times $\gtrsim$ 0.5 d is no better than $\sim$ 0.1 the intensity of the Crab Nebula (approximately the maximum intensity of Her X-1), so that considerations, such as occultation by the Earth, make the average sensitivity of the experiment even poorer. During the total duration of observation reported here (October 16, 1974 through April 16, 1976), a total of 803 data accumulations of $\leqslant$ 0.5 d, for which Her X-1 was not either outside or at the very edge of the experiment field-of-view, have been obtained. Only 24 of these allow intensity definition at the 2σ level (corresponding to $\sim$ 3σ detection of a positive signal above background). Of these 24 measurements, all but one were within the 'on-time' of the 35-d cycle, and a full 1/3 were obtained from a single cycle during which the viewing constraints were near-optimum. It is important to note that the same maximum source intensity produces, for average viewing conditions, a < 1σ signal.

We can, however, take advantage of the fact that we can differentiate between null statistical results under optimum viewing conditions (corresponding to a knowledge that the source was not at maximum intensity), and those at more unfavourable conditions (for which no such knowledge can be gleaned). All the data were, therefore, normalised to the same viewing constraints, with individual negative measurements of the intensity relative to background initially defined to be zero. These 803 points, excluding measurements centred within 7 per cent of the 1.7-d Her X-1 eclipse minimum, were then folded at various trial periods near 35 d. In Fig. 1, where we have plotted the total variance for a 15-bin light curve compared to the average value (essentially equivalent to a χ^2 test against the hypothesis of a constant source intensity), the 35-d modulation of the normalised record is apparent. Figure 1 alone prescribes a period of 34.7±0.3 d (where the width of the variance peak is consistent with the overall sample duration), in agreement with the 34.88±0.12 d of ref. 6. A comparison of the shape and phase of the on-state portion of the light
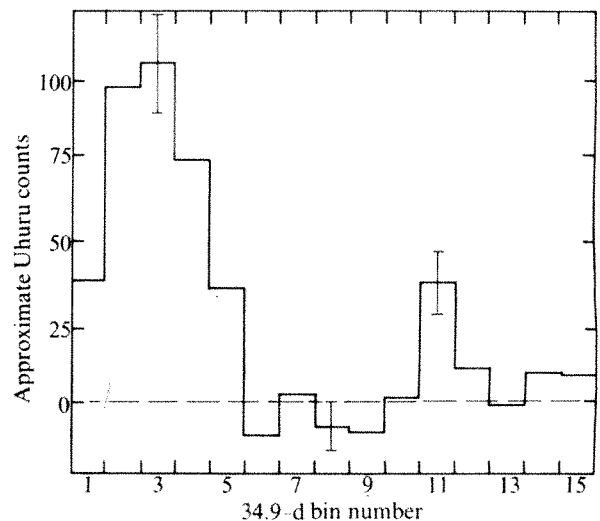
Fig. 2 34.9-d light curve of Her X-1 in 15 bins. The ordinate is normalised to approximate Uhuru counts for comparison with other measurements. The epoch of the start of bin 1 is JD 2,442,442.0.
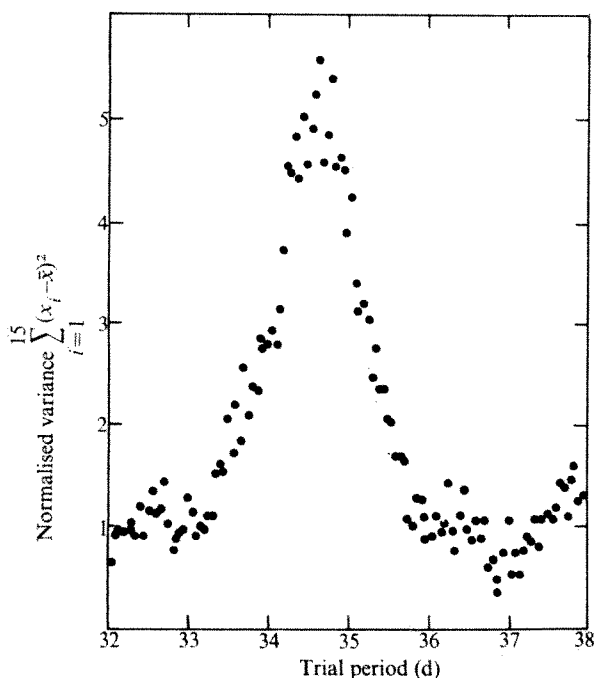
curves obtained here with those of ref. 6 clearly favours a value of 34.9±0.1 d.

The 34.9-d light curve is exhibited in Fig. 2, where the negative intensity measurements are now properly included in the folding procedure (allowing both consistent error estimates and a zero baseline intensity to be obtained). In addition to the characteristic reproduction of the on-state portion of the cycle (bins 1–5), bin 11 contains a > 3σ excess relative to the surrounding off-state intensity. This feature does not arise primarily from the single off-state measurement of high statistical significance, as that datum contributes to bin 12. Jones and Forman have reported that the analagous feature they observed in Uhuru data near January 1, 1972 represented $\sim$ 30% of the maximum intensity, with a duration of $\sim$ 15% of the 35-d cycle (between φ = 0.45 and φ = 0.6 relative to φ = 0 at maximum). The present data agree with the magnitude of the effect, but differ somewhat in phase and duration. In the Ariel V data, the feature is centred at φ = 0.55±0.05 relative to maximum, with an apparent duration of $\leqslant$ 10% (at an intensity $\gtrsim$ 1/4 of maximum). The agreement in magnitude of the present feature with that obtained from Uhuru, coupled with the fact that it is never statistically detectable in an individual measurement, strongly implies that it occurs regularly in every cycle at $\sim$ 1/3 of maximum.

Jones and Forman have attempted to refine model disk geometry parameters on the basis of the Uhuru data, but the present measurements are obviously too crude to revise sensibly their estimates. Clearly, a high-sensitivity measurement of this portion of the light curve is required, and our results would suggest that this experiment is possible during any Her X-1 cycle.

Fig. 1 Variance for 15-bin light curves as a function of trial period. The ordinate is normalised to the mean for trial periods 32.0–33.0 d and 36.0–38.0 d.

S. S. HOLT
E. A. BOLDT
L. J. KALUZIENSKI
P. J. SERLEMITSOS
J. H. SWANK

Laboratory for High Energy Astrophysics,
NASA/Goddard Space Flight Center,
Greenbelt, Maryland 20771

1 Cooke, B. A., and Page, C. G., *Nature*, **256**, 712 (1975).
2 Jones, C., and Forman, W., *Astrophys. J. Lett.* (in the press).
3 Roberts, W. J., *Astrophys. J.*, **187**, 575 (1975).
4 Gerend, D., and Boynton, P. E., *Astrophys. J.* (in the press).
5 Holt, S. S., *Astrophys. Space Sci.*, **42**, 123 (1976).
6 Giacconi, R., *et al.*, *Astrophys. J.*, **184**, 227 (1973).

Microwave spectral lines from interstellar dust

SPECTRAL features at 3.07, 10, 20 and 33 μm have proved invaluable in the chemical characterisation of interstellar dust[1,2]. As spectral observations are extended into the far infrared, it is likely that other features attributable to the absorption and emission of radiation by dust will be observed. To date there has been no discussion in the literature on the possibility of discrete spectral features attributable to dust at even longer wavelengths. The purpose of this note is to suggest that relatively sharp spectral lines may arise from neutral or ionised atomic species in interstellar dust at mm and cm wavelengths. The detection of these spectral features would provide definitive information on the chemical composition of interstellar dust.

The 21-cm line of atomic H arises from the $F = 0 \longleftrightarrow F = 1$ hyperfine transition of ground state H atoms. H atoms can be found in many different types of solid or on solid surfaces (Table 1) when these solids are subjected to nuclear irradiation. The $F = 0 \to F = 1$ transition for H atoms in these systems occurs at the frequencies shown in Table 1. The list has been chosen as representative of solids of astrophysical interest, but it is by no means complete. It serves only to show that H atoms can be stabilised in many types of solid astrophysical interest. H atoms in these environments will absorb or emit radiation at frequencies within $\pm \sim 30$ MHz of 1,420 MHz. The width of the hyperfine line as measured in ESR experiments is typically 0.3–3 MHz. This width arises primarily through spin-lattice coupling[3] for temperatures $\lesssim 50$ K and would be smaller in an astrophysical environment where coupling to lattice acoustic phonons will be diminished, because of the small size of interstellar grains. This effect, which has been seen[4] in laboratory ESR spectra of powders, arises because spin-lattice relaxation at low temperatures involves the creation

Table 1 Frequency of H atom $F = 0 \to F = 1$ hyperfine transition in various solids

System	ν (MHz)	Δν (MHz)	T(K)	Ref.
H in H_2	1,417.11	1	4.2	7
H in silica	1,409.1	1.4–2.8	78	8
H in crystal quartz	1,453.1 ±0.1	< 0.3	78	8
H on silica surface	1,409.3	< 3	77	9
H on silica gel surface	1,411.5		77	10
	1,417	≃ 3	77	
H in H_2O	1,400	≃ 20	4	11
H on SiO_2: Al_2O_3 surface	1,416	≃ 3		12
H on NH_4Y and HY zeolites	1,400	≃ 9		13
	1,406	≃ 3.5		
H in CH_4	1,412	≃45	4.2	14

of acoustic phonons whose energies are close to that of the hyperfine transition. In small particles a cutoff of the acoustic phonon spectrum occurs at $v_c \simeq v_s/2a$ where v_s is the speed of sound in the solid[5] ($\simeq 10^5$ cm s^{-1}) and a is the particle radius. The value of $v_c \simeq 5 \times 10^9$ Hz, for $a = 10^{-5}$ cm; and $\simeq 5 \times 10^{11}$ Hz for $a = 10^{-7}$ cm. Both these frequencies are $> 1,420$ MHz, and thus single-phonon spin-lattice relaxation would not be important for H atoms in interstellar grains. Broadening due to spin–spin interactions, which is proportional to spin concentration[3], could be important in larger particles but would not be expected to be significant in small particles containing only a few H atoms. We conclude that $\Delta v \simeq 0.3$–3 MHz is likely to be an upper limit to the width of the 21-cm line for H atoms trapped in interstellar grains.

Table 2 Representative data on zero field transitions of Fe^{3+} and Cr^{3+} ions in various crystals

System	ν(GHz)	Δν(MHz)	Ref.
Fe^{3+} –SBQ*	7.1132±0.0002	10–20	15
	8.8125±0.0002		
Fe^{3+}–NA†	24.1575±0.0005	10–20	15
	35.365±0.002		
Fe^{3+} –Al_2O_3	12.07		16
	19.13		
Fe^{3+} –TiO_2	43.3		17
	81.3		
Fe^{3+} –MgO	1.844		Calculated
Fe^{3+} –CaO	0.5846		Calculated
Fe^{3+} –$CaCO_3$	6.550±0.4		18
	11.03±0.18		
Cr^{3+} –Al_2O_3	11.5		19
Cr^{3+} –TiO_2	43.3		20
Cr^{3+} –$MgAl_2O_4$	29.7		21

*SBQ, Synthetic brown quartz.
†NA, Natural amethyst

H atoms could be formed in interstellar grains through the action of cosmic rays or photons from the interstellar ultraviolet radiation field. Alternatively, H atoms from the ambient gas may collide with and stick to grains. To estimate the column density required to yield measurable absorption one can use the relationship between n_HL, the number of H atoms in grains per cm^2, and the optical depth τ_v given by Kerr[6].

$$n_HL = 3.88 \times 10^{14} T_s \tau_v \Delta v$$

where T_s is the excitation temperature, and Δv is the linewidth in Hz. Taking $T_s = 10$ K as typical of the temperature of an interstellar grain, $\tau_v = 10^{-2}$ and $\Delta v = 10^5$ Hz, one has $n_HL = 3.9 \times 10^{18}$ cm^{-2}. Then, n_gL, the column density of grains $= 3.9 \times 10^{18}/n$ where n is the number of H atoms per grain. For a 1% concentration of H atoms in grains one obtains $n_HL \simeq 10^{24}$ cm^{-2} independent of the choice of a when the usual mass ratio of gas to dust is used and the dust is assumed to consist of a material with density $\simeq 2$ g cm^{-3} made up of the heavier elements. Though column densities as great as this are not found in typical interstellar clouds, they may be found in protostar clouds or regions of strong infrared absorption or emission[1].

Zero-field splitting[3] of the ground state of many transition metal ions in solids occurs when the crystal field partially removes the M_J degeneracy of states with $J > 1/2$. As an example, the ground state of Fe^{3+} is $^6S_{5/2}$ in the free ion. In a crystal with other than cubic symmetry, this state splits into 3 doubly degenerate levels with $m_J = \pm 1/2, \pm 3/2$ and $\pm 5/2$. Magnetic dipole transitions with $\Delta m_J = 0, \pm 1$ are permitted between these levels. Table 2 shows that these transitions occur in the frequency range between 6 and 80 GHz. In a crystal field with cubic symmetry (MgO, CaO) the ground state of Fe^{3+} splits into only 2 levels, and in that case a single spectral line is ob-

served. Taking a spectral line at $v = 100$ GHz with $\Delta v = 10$ MHz as typical of these transitions one obtains

$$n_{Fe}L \simeq 2 \times 10^9 \, T_s \tau_v \Delta v$$

where

$$T_s = 10 \text{ K}, \, \tau_v = 10^{-2}, \text{ and } n_gL \simeq 2 \times 10^{15}/n$$

where n is the number of Fe^{3+} ions per grain. The corresponding hydrogen column density is $N_HL = 10^{21} \text{ cm}^{-2}$ when one assumes a Fe^{3+} concentration in dust of 1%. Our conclusion is that transitions between the zero-field levels of Fe^{3+} or other transition metal ions in interstellar dust may be observable in absorption in typical interstellar dust clouds.

Though we have emphasised the possibility of detecting absorption features attributable to dust in the mm and cm regions it may be more practical to search for emission features. Grains should be relatively transparent in this region, because of the absence of optical phonon modes at such low frequencies. This implies that most of the energy radiated by grains at mm and cm wavelengths will be radiated within these spectral lines. In view of the optical pumping of interstellar grains that occurs through absorption of short-wavelength photons from the interstellar radiation field, non-thermal emission at the frequencies of these discrete features is a distinct possibility. Microwave lines at 81.541 GHz and 43.122 GHz, attributed to emission from rotational levels of the SiO $v = 1$ state[22], are both within 0.2 GHz of the zero-field lines from Fe^{3+}–TiO_2 (Table 2).

This research was supported by a grant from the NRC of Canada.

W. W. DULEY

Physics Department,
York University,
Toronto, Ontario, Canada

1 Gillett, F. C., and Forrest, W. J., *Astrophys. J.*, 179, 483–491 (1973).
2 Hagen, W., Simon, T., and Dyck, H. M., *Astrophys. J. Lett.*, 201, L81–84 (1975).
3 Carrington, A., and McLachlan, A. D., *Introduction to Magnetic Resonance* (Harper and Row, New York, 1967).
4 Honig, A., *Paramagnetic Resonance* (edit. by Low, W.), 439–442 (Academic, New York, 1963).
5 Duley, W., W., *Nature, phys. Sci.*, 244, 57–59 (1973).
6 Kerr, F. J., *Nebulae and Interstellar Matter* (edit. by Middlehurst, B. M., and Aller, L. H.), 575–622 (University of Chicago Press, Chicago, 1968).
7 Jen, C. K., Foner, S. N., Cochran, E. L., and Bowers, V. A., *Phys. Rev.*, 112, 1169–1182 (1958).
8 Weeks, R. A., and Abraham, M., *J. chem. Phys.*, 42, 68–71 (1965).
9 Ogawa, S., and Fessenden, R. W., *J. chem. Phys.*, 41, 1516–1517 (1964).
10 Atkins, P. W., and Symons, M. C. R., *The Structure of Inorganic Radicals* (Elsevier, New York, 1967).
11 Piette, L. H., Rempel, R. C., Weaver, H. E., and Flournoy, J. M., *J. chem. Phys.*, 30, 1623–1624 (1959).
12 Vedrine, J., Dalmai, G., and Imelik, B., *J. Chim. phys.*, 65, 1780–1793 (1968).
13 Abou-Kais, A., Vedrine, J., Massadrier, J., and Dalmai, G., *J. Chim. phys.*, 69, 570–573 (1972).
14 Rexroad, H. N., and Gordy, W., *Phys. Rev.*, 125, 242–247 (1962).
15 Cook, A. R., and Matarrese, L. M., *J. chem. Phys.*, 50, 2361–2364 (1969).
16 King, J. E., and Terhune, R. W., *J. appl. Phys.*, 30, 1844–1845 (1959).
17 Carter, D. L., *J. appl. Phys.*, 32, 2541–2542 (1961).
18 Wakabayashi, J., *J. chem. Phys.*, 38, 1910–1912 (1963).
19 Kikuchi, C., Lambe, J., Makhov, G., and Terhune, R. W., *J. appl. phys.*, 30, 1061–1067 (1959).
20 Gerritsen, H. J., Harrison, S. E., and Lewis, H. R., *J. appl. Phys.*, 31, 1566–1571 (1960).
21 Stahl-Brada, R., and Low, W., *Phys. Rev.*, 116, 561–564 (1959).
22 Lovas, F. J., *Astrophys. J.*, 193, 265–272 (1974).

Force-free magnetic fields in the fluid interiors of neutron stars

I SHOW here that, in equilibrium, the magnetic field in the fluid interior of a neutron star satisfies the condition $\mathbf{B} \times \nabla \times \mathbf{B} = 0$, which is the usual condition for a field to be force free[1]. The significance of this result is that it places a great theoretical restriction in the possible structure of such fields. Restrictions such as this are important because it is still very difficult to deduce anything from observation about the magnetic fields inside neutron stars.

In the region of the star considered here, the density lies between approximately $2 \times 10^{14} \text{ g cm}^{-3}$ and a few times this

value. The material in that region is a fluid composed of neutrons, protons and electrons[2,3]. It is a simple matter to extend the proof to include the effects of muons and hyperons which may appear at higher densities[2]. Although calculations indicate that the neutrons are superfluid and the protons are superconducting at the temperatures expected inside neutron stars[2,3], the proof of the force-free nature of the field does not depend on these assumptions. Newtonian gravitational theory is used here; the general relativistic proof is essentially similar. The proof may also be extended to rotating neutron stars by including centrifugal terms in the chemical potentials.

The crucial part of the argument is that the dynamical β-equilibrium between the neutrons, protons and electrons requires that

$$\mu_e + \mu_p = \mu_n \qquad (1)$$

where the μs are chemical potentials. These potentials include rest mass energies, Fermi energies and gravitational potential energies per particle. The potentials also depend to a very slight degree on the magnetic field strength[1]. The equilibrium of the neutrons implies that their chemical potential is constant throughout the region

$$0 = -\nabla \mu_n \qquad (2)$$

This equation expresses the hydrostatic equilibrium of the neutrons.

But $-\nabla \mu_e$ and $-\nabla \mu_p$ do not vanish in the presence of a magnetic field, unlike $-\nabla \mu_n$, because both the electrons and protons experience Lorentz forces. If the electron and proton analogues to equation (2) are added together, the result is an equation expressing the magnetohydrostatic equilibrium of the electron–proton plasma

$$0 = -\nabla(\mu_e + \mu_p) + \frac{1}{cn_p} \mathbf{J} \times \mathbf{B} \qquad (3)$$

In deriving the equation, which holds in the centre-of-mass frame of the plasma, charge neutrality ($n_e = n_p$) has been used. It follows directly from equations (1), (2) and (3) that

$$\mathbf{J} \times \mathbf{B} = 0 \qquad (4)$$

where the conduction current density is denoted $\mathbf{J}$. No distinction is made here between the magnetic field $\mathbf{H}$ and the magnetic induction $\mathbf{B}$ because the magnetic permeability is effectively unity. At the densities considered here, the spacing between Landau levels is very much smaller than the Fermi energies (V. Canuto, unpublished). In addition, for a wide range of reasonable temperatures and field strengths, the spacing is smaller than, or comparable to, kT. As a result of these conditions the de Haas–van Alphen oscillations in the magnetic permeability can be neglected.

Because of the previously noted slight dependence of the chemical potentials on the magnetic field strength, there is a small force on the plasma attributable to the magnetic field. But for the reasons mentioned with regard to the magnetic permeability, this force is relatively insignificant (ref. 4, and I.E., unpublished).

It follows from equation (4) that $\mathbf{B}$ satisfies

$$\mathbf{B} \times \nabla \times \mathbf{B} = 0 \qquad (5)$$

which is known as the 'force-free' condition[1].

It should be noted that this condition on $\mathbf{B}$ is true in spite of a small force on the plasma, resulting from the field. The reason the magnetic fields in many other astrophysical plasmas are believed to be force free or nearly so is the smallness of the plasma pressure compared to the magnetic pressure[1]; the

reason for the force-free nature of the field in the neutron star fluid interior is entirely different. The essential point here is that in β-equilibrium, the electrons and protons 'communicate' with each other through the neutrons, causing $-\nabla(\mu_e + \mu_p)$ to vanish, in turn causing $\mathbf{J} \times \mathbf{B}$ to vanish. Furthermore, since the standard stability analyses of force-free fields do not take rotation into account[1], they are not applicable to many neutron stars, in which rotation may profoundly affect the interior plasma dynamics (refs 5 and 6, and I.E., unpublished).

As one simple example of the theoretical use to which equation (5) can be put, note that it can eliminate from consideration any non-singular, purely toroidal field, such as that used by Vandakurov[7].

I acknowledge useful conversations with Professors D. Q. Lamb, C. Pethick and M. H. L. Pryce. This research was supported by the NSF.

IAN EASSON

Department of Physics,
University of Illinois at Urbana-Champaign,
Urbana, Illinois 61801

1 Korchak, A. A., and Moledensky, M. M., *Commun. Astrophys. Space Phys.*, **6**, 81–85 (1975).
2 Baym, G., and Pethick, C., *A. Rev. Nucl. Sci.*, **25**, 27–77 (1975).
3 Baym, G., Pethick, C., and Pines, D., *Nature*, **224**, 673–674 (1969).
4 Shul'man, G. A., *Sov. Phys. J.*, **17**, 1355–1358 (1976).
5 Hide, R., *Nature*, **229**, 114–115 (1971).
6 Easson, I., thesis, Univ. British Columbia, (1975).
7 Vandakurov, Yu., *Astrophys. Lett.*, **5**, 267–269 (1970).

A rare event in the stratosphere

PERTINENT to the question of the existence of haze and water vapour in the stratosphere[1,2] was the detection of a nacreous cloud at 25 km over New Mexico on June 20, 1974. This occurred just before the volcanic dust incursion from Fuego in November, 1974 (ref. 3), during a high altitude balloon experiment in which a large polar nephelometer was used to measure the optical angular-scattering characteristics of atmospheric aerosols[4].

The balloon launch was made at 0748 MST from Holloman AFB, New Mexico (32°N, 107°W) into a very clear sky. The nephelometer reached an altitude of 26.6 km and payload ground-impact was 7 h later near Silver City, New Mexico, 210 km W of the launch site. Easterly winds of ~ 20 knots carried the experimental package on a trajectory perpendicular to the San Andres mountain range. The minimum ambient of −75 °C was reached at 16.8 km, and the balloon was still ascending when the very strong signal was encountered on the lee side of the mountains. The intense signals continued to the apogee of the flight at 26.6 km without diminution, thereby implying a greater vertical extent for the increased aerosol particle concentrations. The balloon was allowed to float at this maximum altitude for 15 min, evidently immersed in the cloud. Moreover, the signals continued for a period of 50 min, during which time the balloon had travelled nearly 50 km with respect to the ground, thus making the latter distance the minimum horizontal extent of the cloud. On the descent of the balloon, the scattering ceased as dramatically as it had developed on ascent.

Figure 1 shows a typical altitude profile of the volume scattering function (the absolute intensity) measured at a scattering angle of 15° and spectral wavelength of 0.475 μm. Four other photometers, at angles of 30°, 50°, 100° and 150° gave a similar intensity enhancement, although proportionately reduced by the characteristic forward trend of scattering from large particles. The high intensity at ground level (1.29 km above sea level) is caused by exhaust smoke from the balloon launch crane, while that from 3–7 km results from increased amounts of aerosol particles in the convective mixing layer. The stratospheric haze is graphically illustrated by the large variations in intensity as a result of altitudinal inhomogeneity in the particle concentration. The payload also apparently descended through a Cirrus cloud at ~ 10.6 km.

The unusual occurrence of a large haze cloud at 25 km—though it was not evident from the ground—was fortuitous, and is rarely observed in summer at mid-latitudes. Stanford and Davis[5] have made a review of stratospheric cloud reports for the period 1870–1972, and of some 183 sightings only 3 were reported in the USA. Of these three, two events occurred near the San Andres mountains and the remaining one in northern Arizona. This suggests that the mountains of New Mexico may cause the development of stratospheric clouds; however, their visual observation is
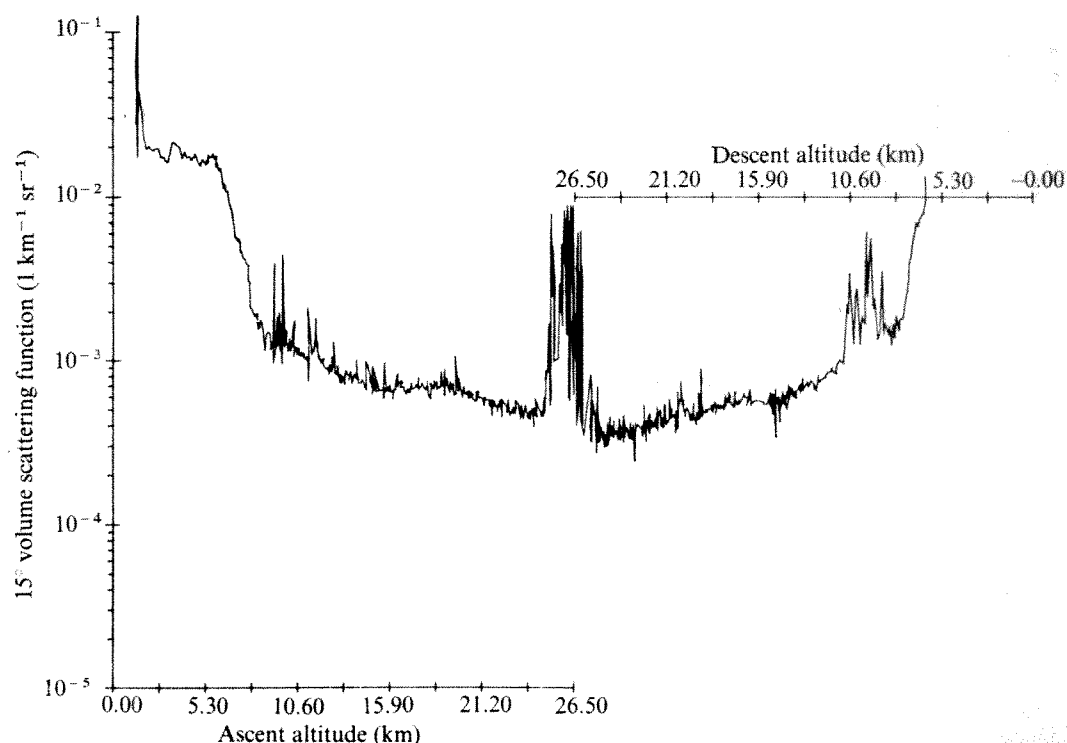
Fig. 1 The 15° volume scattering function against altitude at a wavelength of 0.475 μm.

difficult because of their high altitude and tenuous structure. Nonetheless, the rarity of sightings may result in part simply from a lack of observation facilities (or an observer) rather than infrequency of occurrence of such events.

Finally, this cloud provides experimental evidence relevant to Scorer's[2] suggestion, that, even in the absence of volcanic dust, air–mass interchange mechanisms across the tropopause can result in extensive haze formations at high altitudes. Moreover, observations made *in situ* logically provide more reliable information on the stratospheric optical environment, since atmospheric dynamics in the troposphere not only produce sundry effects in the stratosphere but complicate, in particular, remote measurements, since scattering properties along the ground-to-stratosphere light transmission path must be considered.

FRANK W. GIBSON

Air Force Geophysics Laboratory,
Hanscom AFB, Bedford,
Massachusetts 01731

1 Sissenwine, N., Grantham, D. D., and Samela, H. A., *J. atmos. Sci.*, **25**, 1529 (1968).
2 Scorer, R. S., *Nature*, **258**, 134 (1975).
3 Meinel, A. B., and Meinel, M. B., *Science*, **188**, 477 (1975).
4 Gibson, F. W., *J. opt. Soc. Am.*, **65**, 1174 (1975).
5 Stanford, J. L., and Davis, J. S., *Bull. Am. met. Soc.*, **55**, 213 (1974).

Gas phase chemical kinetics of sodium in the upper atmosphere

THE existence of a layer of atomic sodium in the upper atmosphere near 90 km has been clearly demonstrated by analyses of daytime and twilight airglow emissions of D-line radiation at 5,890 and 5,896 Å (refs 1 and 2) as well as by direct laser backscatter measurements (see, for example ref. 3). Following Chapman[4] chemical reactions involving atomic sodium and atmospheric oxygen species have been invoked to explain the chemical dynamics of the sodium layer and the nightglow emission of D-line radiation. In addition, Chapman has also proposed that similar reactions involving newly ablated sodium are responsible for the occurrence of long enduring visible meteor trails[5]. Unfortunately, previous quantitative attempts to model sodium-induced upper atmospheric chemiluminescence have suffered from rather poor estimates of the appropriate reaction rate constants. Here we provide more reasonable estimates of rate constants and show how these new kinetics parameters can dramatically alter calculations of chemiluminescent efficiency.

The sodium–oxygen atmospheric reaction sequence is:

$$Na + O_3 \xrightarrow{k_1} NaO + O_2 \qquad (1)$$

$$NaO + O \xrightarrow{fk_2 \text{ and } (1-f)k_2} \begin{cases} Na\,(^2P) + O_2 & (2a) \\ Na\,(^2S) + O_2 & (2b) \end{cases}$$

$$NaO + O_3 \xrightarrow{k_{3a} \text{ and } k_{3b}} \begin{cases} NaO_2 + O_2 & (3a) \\ Na + 2O_2 & (3b) \end{cases}$$

$$Na + O_2 + M \xrightarrow{k_4} NaO_2 + M \qquad (4)$$

$$Na + O + M \xrightarrow{k_5} NaO + M \qquad (5)$$

$$NaO_2 + O \xrightarrow{k_6} NaO + O_2 \qquad (6)$$

$$Na(^2P) \xrightarrow{k_7} Na\,(^2S) + h\nu\,(D_1, D_2) \qquad (7)$$
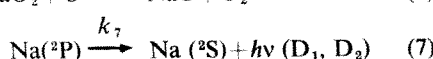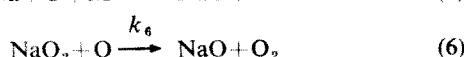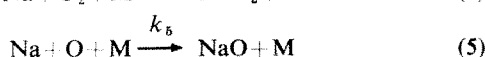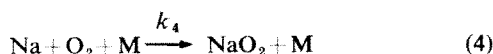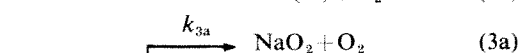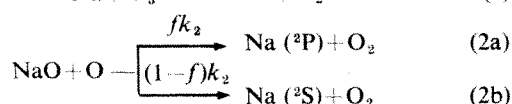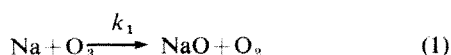
where all species are in their electronic ground states unless otherwise noted. The combination of reactions 1, 2a and 7 in

Table 1 Ionisation data for sodium and oxygen species

Species	IP (eV)	EA (eV)	Reference
Na	5.14		19
NaO	6.5		12
O_3		1.99	20
O		1.46	21

the presence of O and O_3 in the upper atmosphere form a chemiluminescent chain and may explain both sodium nightglow and meteor trail emissions. Reaction 7 is sufficiently fast ($k_7 = 1/\tau_{rad} = 6.3 \times 10^{-9}$ s^{-1}) that all Na (^2P) created in reaction 2a radiates before quenching at upper atmospheric densities. Recent observations of the nightglow D-line profiles are consistent with this sodium oxide chemiluminescent cycle[7,8].

Several recent attempts to model sodium–oxygen chemistry in the upper atmospheric sodium layer[1,2,9] and in meteor trail wakes[10] quantitatively have been published. These efforts have been severely hampered by the fact that experimental rate constants are only available for reactions 5 and 7. To overcome this handicap, most authors have adopted experimental

Table 2 Calculated reactive cross sections, rate constants and exothermicities for sodium–oxygen reactions

Reaction	σ_R(cm^2)	k_R(cm^3 s^{-1})	ΔH_R^0 (kcalorie mol^{-1})
1	6.5×10^{-15}	3.3×10^{-10}	-34.8 ± 4
2	2.6×10^{-15}	1.6×10^{-10}	$a - 9.2 \pm 4$ $b - 57.7 \pm 4$
3	3.2×10^{-15}	1.4×10^{-10}	$a - 97.2 \pm 7$ $b - 150.2 \pm 7$

rate constants for reactions of either atomic H or N and OH or NO in place of atomic Na and NaO. Values of the order of 10^{-11}–10^{-13} cm^3 s^{-1} for k_1 and 10^{-11} cm^3 s^{-1} for $k_{(2a+2b)}$ are typical in these studies.

Unfortunately, these rate constant estimates are very likely to be substantially low. Exothermic bimolecular processes, like reaction (1) between low ionisation potential species such as alkali atoms and electrophilic molecules such as O_3 are known to occur by an electron jump mechanism[11]. Since the measured ionisation potential of NaO is only 6.5 ± 0.7 eV (ref. 12), it is quite likely that reactions (2) and (3) proceed by electron jump mechanisms as well.

Relatively accurate reactive cross sections for exothermic electron jump reactions can be calculated by determining the radius at which the Coulombic attraction of the ion pair formed by the electron jump, equals the energy deficit between the ionisation potential of the metal containing reactant, IP, and the electron affinity of the electrophilic reactant, EA

$$e^2/r_x = \text{IP} - \text{EA} \qquad (8)$$

The calculated reactive cross section is merely[11,13]

$$\sigma_R \simeq \pi r_x^2 \qquad (9)$$

Using the thermochemical parameters shown in Table 1, equations (8) and (9) can be used to estimate overall reactive cross sections for reactions (1), (2) and (3). Rate constants for these reactions can also be estimated by multiplying the reactive cross sections by the most probable relative velocities for reactant pairs at typical upper atmospheric temperatures.

Table 2 shows reactive cross sections and rate constants for an ambient temperature of 185 K calculated by the procedures outlined above. Table 2 also lists the exothermicities of reactions (1), (2) and (3) calculated using the Na–O bond strength of 60.3 ± 4 kcalorie mol^{-1} measured by Hildenbrand and Murad[12], the Na–O_2 bond strength of 65 ± 3 kcalorie mol^{-1} measured by McEwan and Phillips[14] as well as O–O and O–O_2 bond strengths

of 118.0 and 25.5 kcalorie mol^{-1}, respectively. Furthermore, the Na–O, Na–O$_2$ and O–O bond energies can be combined to yield an estimate of the NaO–O bond strength of 122.7 ± 7 kcalorie mol^{-1}.

It should be noted that the rate constants for reactions (1) and (2), listed in Table 2 are significantly larger than the analogous rate constants for hydrogen or nitrogen species used in previous modelling efforts[1,2,9,10]. As an example of the impact these larger rate constants can have on an atmospheric modelling problem, the upper limit of Na D-line emissions from a ~ 30-g meteoroid with a velocity of 40 km s^{-1} will be calculated below. Following Baggaley[10] such a meteoroid is assumed to deposit a line density of atomic Na of 2×10^{15} cm^{-1}. Atmospheric abundances and common molecular diffusion coefficients, D, are also adopted from Baggaley's work and are shown in Table 3. The values adopted for O and O$_3$ number densities are reasonably consistent with recently published measurements[15-17].

The reaction scheme used in the calculation of meteor trail D-line luminosity consists of reactions (1), (2), (3a), (4), (5) and (7) listed above. Calculations were performed with the branching fraction, f, in reaction 2 set equal to 1 to yield an upper limit to the D-line emission. Reaction (6) is deleted since the bond values listed above indicate that it is close to thermoneutral and is probably slow at mesospheric temperatures. The exclusion of reaction 6 makes NaO$_2$ a chain-terminating species in night-time conditions. Only the 'a' channel of reaction (3) was considered. Inclusion of (3b) would have hindered NaO$_2$ formation and yielded a higher photon production term.

The rate constants for reactions (1), (2) and (3) were used as calculated and listed in Table 2. The rate for reaction (4) has been measured by Carabetta and Kaskan[18] in flame studies and was found to equal 8.2×10^{-34} cm^6 s^{-1} between 1,420 and 1,600 K, independent of temperature. This rate was adopted for both three-body processes (4) and (5). Reaction (7) is essentially instantaneous on the time scale of mesospheric collisions.

The set of equations describing the interaction of meteor deposited Na with the atmospheric oxygen species may be written

$$\frac{d}{dt}[Na] = D\nabla^2[Na] - k_1[Na][O_3] + k_2[NaO][O]$$
$$- k_4[Na][O_2][M] - k_5[Na][O][M]$$

$$\frac{d}{dt}[NaO] = D\nabla^2[NaO] + k_1[Na][O_3] - k_2[NaO][O]$$
$$- k_3^a[NaO][O_3] + k_5[Na][O][M]$$

$$\frac{d}{dt}[NaO_2] = D\nabla^2[NaO_2] + k_3[NaO][O_3]$$
$$+ k_4[Na][O_2][M]$$

$$\frac{d}{dt}[O_3] = D\nabla^2[O_3] - k_1[Na][O_3] - k_3[NaO][O_3]$$

$$\frac{d}{dt}[O] = D\nabla^2[O] - k_2[NaO][O] - k_5[Na][O][M]$$

$$I = f\Sigma k_2[NaO][O] \text{ photons s}^{-1}\text{ cm}^{-1}$$

where ∇^2 is the Laplacian operator in cylindrical coordinates. Reaction with meteor-produced atomic oxygen is not included in this model since Baggaley's treatment indicated that this
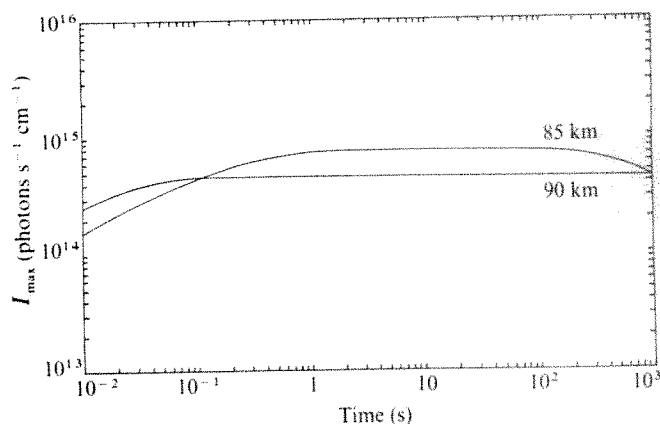
Fig. 1 Maximum D-line photon emission rates at 85 and 90 km.

effect does not persist past the first 10 s[10]. An initial meteor train radius of 100 cm was assumed.

Solution of the set of equations listed above was performed using a modified Crank–Nicholson finite difference approach. Solutions of I_m, the maximum photoemissions ($f = 1$), for both the 85- and 90-km cases are shown in Fig. 1. The maximum photon emissions found in this study are 4.6×10^{14} and 7.5×10^{14} photons s^{-1} cm^{-1} at 90 and 85 km, respectively. These compare with peak values of 8×10^{10} and 3×10^{11} calculated by Baggaley using a k_1 of 10^{-13} and a k_2 of 10^{-11} cm^3 s^{-1}. The slow decay of photon intensity after 100 s seen in the 85-km solution is caused by chain-terminating formation of NaO$_2$ due to the deletion of reaction 6.

The large increases (> 10^3) in calculated maximum D-line photon emissions due to the use of more realistic sodium–oxygen reaction rates are clearly important in determining if Chapman's proposed mechanism[5] can be responsible for long lived visible meteor trails at high altitudes. Although no definitive statement can be made without a firm knowledge of f, the branching fraction for excited sodium production in reaction (2), the results shown in Fig. 1 clearly lend more support to Chapman's proposal than Baggaley's recently published calculation[10].

Some feeling for the magnitude of f can be obtained by comparing observed nightglow emissions with those predicted by reactions 1–7 using our rate constant estimates. Quantitative precision is difficult because of the wide variations observed in both ambient Na densities and D-line nightglow densities[2,3,7]; however, a value of f between 10^{-1} and 10^{-2} would seem necessary to match observed nightglow intensities.

It should be noted that the rate constants calculated in this work indicate that sodium is an extremely efficient catalytic destroyer of odd oxygen, (O + O$_3$), at the atmospheric altitudes considered. This efficiency may extend to lower altitudes depending on the magnitude of k_6, the photodissociation cross section for NaO$_2$, and the rate of heterogeneous removal of sodium oxides.

C. E. KOLB
J. B. ELGIN

*Aerodyne Research, Incorporated,
Bedford Research Park,
Crosby Drive,
Bedford, Massachusetts 01730*

1 Blamont, J. E. and Donahue, T. M., *J. geophys. Res.*, **69**, 4093–4127 (1964).
2 Hunten, D. M., *Space Sci. Rev.*, **6**, 493–573 (1967).
3 Bowman, M. R., Gibson, A. J., and Sandford, M. C. W., *Nature*, **221**, 456–457 (1969).
4 Chapman, S., *Astrophys. J.*, **90**, 309–316 (1939).

Table 3 Atmospheric model for meteor trail calculations

Altitude (km)	Diffusion coefficient, D (cm^2 s^{-1})	Species number densities (cm^{-3})			
		O$_3$	O	O$_2$	M (Total)
90	1.0×10^5	3×10^8	3×10^{11}	1.3×10^{13}	0.63×10^{14}
85	4.2×10^4	8×10^8	7×10^{10}	3.2×10^{13}	1.52×10^{14}

5 Chapman, S., in *The Aurora and Airglow* (edit. by Armstrong, E. B., and Dalgarno A.), 204 (Pergamon, Oxford, 1955).
6 Wiese, W. L., Smith, M. W., and Miles, B. M., *Atomic Transition Probabilities*, II, (NSRDS-NBS 22, 1969).
7 Sipler, D. P., and Biondi, M. A., *Geophys. Res. Lett.*, 2, 106–108 (1975).
8 Hernandez, G., *Geophys. Res. Lett.*, 2, 103–105 (1975).
9 Brown, T. L., *Chem. Rev.*, 73, 645–667 (1973).
10 Baggaley, W. J., *Nature*, 257, 567–568 (1975).
11 Herschbach, D. R., *Adv. Chem. Phys.*, 10, 319–393 (1966).
12 Hildenbrand, D. L., and Murad, E., *J. chem. Phys.*, 53, 3403–3408 (1970).
13 Levine, R. D., and Bernstein, R. B., *Molecular Reaction Dynamics*, 86–87 (Oxford University Press, Oxford, 1974).
14 McEwan, M. J., and Phillips, L. F., *Trans. Faraday Soc.*, 62, 1717–1720 (1966).
15 Dickinson, P. H. G., Bolden, R. C., and Young, R. A., *Nature*, 252, 289–291 (1974).
16 Trinks, H., *Geophys. Res. Lett.*, 2, 99–102 (1975).
17 Noxon, J. F., *J. geophys. Res.*, 80, 1370–1373 (1975).
18 Carabetta, R., and Kaskan, W. E., *J. phys. Chem.*, 72, 2483–2489 (1968).
19 Franklin, J. L., *et al.*, *Ionization Potentials, Appearance Potentials, and Heats of Formation of Gaseous Positive Ions*, 161 (NSRDS-NBS 26, 1969).
20 Wong, S. F., Vorburger, T. V., and Woo, S. B., *Phys. Rev. A*, 5, 2598–2604 (1972).
21 Hotop, H., Bennett, R. A., and Lineberger, W. C., *J. chem. Phys.*, 58, 2373–2378 (1973).

Late Weichselian geomagnetic 'reversal' as a possible example of the reinforcement syndrome

A LATE Weichselian geomagnetic reversal has been reported[1,2] from Sweden. Mörner and Lanser[3] believed the Swedish reversed palaeomagnetic record to be a "real dipole magnetic reversal of greatest significance for global correlations" and to have been established and well dated in Canada, the North Atlantic and New Zealand and tentatively recognised in Czechoslovakia, the Northern Pacific, the Gulf of Mexico, Japan and France. Noël[4] has described further palaeomagnetic results from two Swedish varved-clay cores as accurately tracing the annual behaviour of the geomagnetic event and delimiting its initiation and termination. Recognition of the event in lavas and deep sea, continental shelf, varved-clay and lake sediments has been cited as confirming the reality of the event. That previous research, particularly in Sweden, has concentrated on palaeomagnetic measurements at numerous, different late Weichselian localities. Because of the difficulties in correlating between localities, interpreting palaeomagnetic records in coarse grained sediments and deciphering coring and sedimentological complexities in single cores, we have instead

studied one carefully selected site in detail. We have come to the conclusion that Swedish late Weichselian palaeomagnetic data do not record a reversal or excursion of the geomagnetic field, but demonstrate that the Swedish geomagnetic field has been of normal polarity since 13,000 yr BP. The Swedish late Weichselian reversal had thus become an example of Watkins'[5] "reinforcement syndrome".

We have analysed 408 subsamples from two main cores collected at Björkeröds Mosse on Mt Kullen in north-western Scania in southern Sweden (56.2°N, 12.6°E). Björkeröds Mosse was chosen because it lies in the area considered to be the earliest ice-free part of Sweden[6]; its stratigraphy, vegetational and soil developments have been studied in detail[7,8]; it had a high rate of deposition of organic sediments; and its chronology is well known from both pollen and radiocarbon dating[7]. The sediments from Björkeröds Mosse span the supposed age of the late Weichselian geomagnetic reversal of 12,400 or 12,100 yr BP. Ten sections of 'D'-shaped core, each 1 m in length, were obtained using a Jowsey sampler[9]. The natural remanent magnetisation (NRM) of the 408 subsamples was measured on a 'Digico' magnetometer[10]. Twenty-two pilot samples were partially a.c. demagnetised stepwise in increasing peak fields. A peak alternating field of 100 oersted was selected to eliminate viscous components. The remaining 386 subsamples were cleaned at this field strength. NRM, cleaned declination and inclination, cleaned intensity, and low field reversible susceptibility of core III (1975) are shown in Fig. 1.

We interpret the palaeomagnetic directions as follows. In general, the geomagnetic field close to the time of deposition is recorded by the palaeomagnetic remanence. At a few lithologically distinct horizons, however, the magnetic remanence has been controlled by other processes. The scattered or offset declinations and inclinations in the laminated deposits of layer 1 at depths of 615, 520 and 440 cm (Fig. 1) occur in coarser grained sediments and are due to rough alignment in higher energy deposition environment, rather than to idealised alignment along the ancient magnetic field direction. The single, reversed inclination at a depth of 338 cm (Fig. 1) falls in a narrow dark layer about 2 cm thick. A dark layer at a similar stratigraphic level in Björkeröds Mosse has been reported by Mörner in core B892 to be "well established" as a geomagnetic excursion and to correspond to his Fjards interstadial. The layer

Fig. 1 Data from Björkeröds Mosse core III (1975). ○, NRM; ●, partial demagnetisation at 100 oersted. Crosses indicate end of core sections. Mean declinations were set to zero.
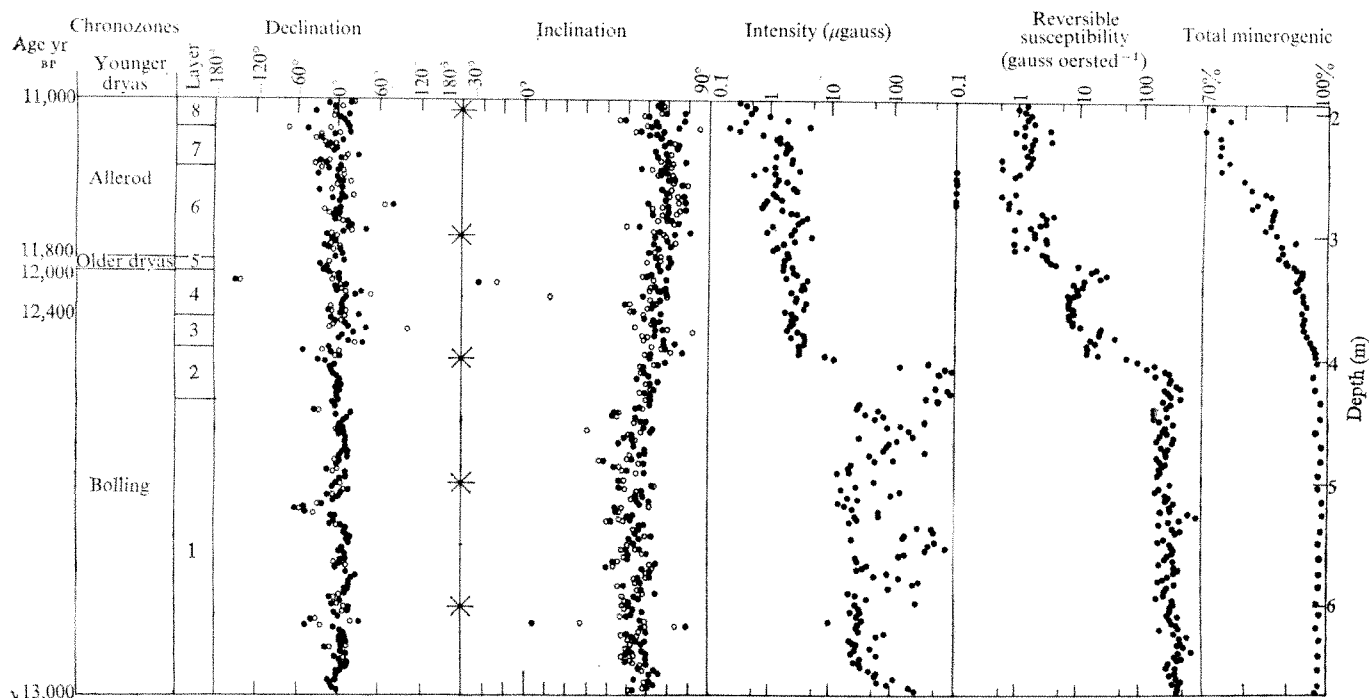
in our core III (1975) is anomalous as, (1) its base is markedly (our results, in preparation) crenelated, (2) it does not occur in two other duplicate cores, and (3) the remanence is unusually stable as well as in a distinct direction. Its pollen spectrum and sediment chemistry correspond to the till which forms the banks of the lake. We prefer to interpret this intermediate palaeo-magnetic direction and Mörner's[11] in core B892 to be due to slumping rather than a true reflection of the geomagnetic field. The remaining great majority of palaeomagnetic directions are all of normal polarity and are taken to indicate that the geo-magnetic field was of normal polarity from 13,000 to 11,000 yr BP in Sweden.

The retention of aberrant magnetic directions and the between-core repeatability (our results, in preparation) of cleaned remanent intensity indicate that these normal polarity late Weichselian palaeomagnetic directions are almost con-temporary with deposition and are unlikely to have been caused by later remagnetisation. Our data suggest previously reported reversed palaeomagnetic directions[1 -4] are not reliable indicators of the ancient geomagnetic field, but have been distorted by mechanical sedimentation processes, slumping or weathering. Alternatively, but less likely, the reversed directions had been dated inaccurately. We suggest the proliferation of unusual palaeomagnetic directions in Scandinavia around 12,000 yr BP is a reflection of changing climatic conditions. In many localities the fluctuations of climate produced sediments of very variable mechanical properties, particularly at times of periglacial activity, which were poor recorders of the ambient magnetic field direction.

Using geomagnetic events in the Swedish late Weichselian is thus of doubtful value. Furthermore, reversals or excursions of the geomagnetic field from other parts of the world around this time should not be dated or correlated with the earlier Swedish results.

Watkins'[12] warnings on "the special problems posed by the search for short events" have unfortunately not been heeded. We endorse his comments[12] and on the basis of our work on the Swedish late Weichselian propose three minimum requirements for the recognition of geomagnetic excursions in recent sedi-ments: (1) both magnetic declination and inclination changes should be repeatable to within 20° in at least two sections or cores at the same site; (2) results should only be taken from macroscopically homogeneous lithologies with an average grain size below 62.5 µm; and (3) magnetic directions should be stable under alternating field cleaning with peak field ≥ 100 oersted.

Application of these criteria would prevent many spurious palaeomagnetic directions being interpreted as evidence for unusual geomagnetic behaviour. It would reduce the large number of excursions and events deduced from Brunhes sediments to a small total comparable to that recognised in igneous material[13]. It would also reduce the difficult dating problems arising from attempts to correlate recent excursions, dispel speculation[11] on very rapid changes in the main dipole field and enable geomagnetic excursions to be considered again as chronological indicators.

R. THOMPSON

Department of Geophysics,
Edinburgh University,
Edinburgh, UK

B. BERGLUND

Laboratory of Quaternary Biology,
Institute of Geology,
Lund University, Sweden

1 Mörner, N-A., Lanser, J. P., and Hospers, J., *Nature phys. Sci.*, **234**, 173-4 (1971).
2 Noël, M., and Tarling, D. H., *Nature*, **253**, 705-707 (1975).
3 Mörner, N-A., and Lanser, J. P., *Nature*, **251**, 408-409 (1974).
4 Noël, M., *Geol. Fören. Stockh. Förh.*, **97**, 357-367 (1975).
5 Watkins, N. D., *Bull. geol. Soc. Am.*, **83**, 551-574 (1972).
6 Nilsson, E., K., *Vetenskapsakad. Handl.* IV, 12:1 (1968).
7 Berglund, B. E., *Geol. Fören. Stockh. Förh.*, **93**, 11-45 (1971).
8 Berglund, B. E., and Malmer, N., *Geol. Fören. Stockh. Förh.*, **93**, 575-589 (1971).
9 Jowsey, P. C., *New Phytol.*, **65**, 245-248 (1966).
10 Molyneux, L., *Geophys. J. R. astr. Soc.*, **24**, 429-434 (1971).
11 Mörner, N-A., *Geol. Fören. Stockh. Förh.*, **97**, 298-301 (1975).
12 Watkins, N. D., *Comments Earth Sci.: Geophys.*, **2**, (2), 36-43 (1971).
13 Dalrymple, G. B., in *Calibration of Hominoid Evolution*, (edit. by Bishop, W. W., and Miller, J. A.), 107-134, (Scottish Academic, Edinburgh 1972).

January-thaw singularity and wave climates along the Eastern coast of the USA

METEOROLOGICAL singularities—recurrent weather phenom-ena at or near specific calendar dates[1]—have been reported throughout the Northern Hemisphere. Although not fully recognised as real phenomena by most meteorologists, singularities continue to merit the attention of some investi-gators. Although the general relationship between wind fields over open oceans and generation of waves by wind are well known, there are no reports in the literature on coastal wave climates and atmospheric singularities. I have observed that extreme frequencies of surf from directions more typical of the summer season occur at the time of the 'January-thaw' meteorological singularity (January 20-23). Here I present observations of wave climates from Florida to Rhode Island at the time of the January thaw.

Slocum[2] recorded an anomalous warm spell during January 20-23, between 1872 and 1921. Subsequently, Wahl[3] demonstrated the existence of the January thaw in New England for the period 1873-1952 and Frederick[4] reported on its extensive geographical distribution between 1927 and 1956. Wahl[3] also noted that at the time of the January thaw the ridge-and-trough positioning over eastern North America differed markedly from its positioning before after-wards; specifically, during the January thaw there is a high pressure cell off the east coast of the USA, and a trough runs through the midwest, whereas immediately following the thaw, the positions of the ridge and trough are reversed. Anticyclonic singularities have been reported for Europe (January 18-24) by Brooks[5] and Japan (January 24) by Maejima[6].

Duquet[7] has attributed the January thaw to an adjust-ment of the planetary circulation occurring in early or mid January. Wahl[8] has noted that the January thaw is most pronounced during low-index Januaries. Significant changes in the general circulation of the atmosphere at the time of the January thaw are well documented. These changes should appear as a singularity in coastal wave-climate statistics.

Surf-height and direction statistics for years falling within a period of general low index[9] were obtained from the Co-operative Surf Observation Program (COSOP) for Point Judith, Rhode Island (1957-1965); Atlantic City, New Jersey (1955-1964); Ocean City, Maryland (1955-1970); Virginia Beach, Virginia (1954-1969); Oak Island, North Carolina (1955-1965); and Hillsboro Inlet, Florida (1955-1965), (Table 1). Daily and 3-d running means of surf heights were examined but did not show any anomalous characteristics at the time of the singularity; however, at the time of the January-thaw singularity, the directions of approaching surf along the Atlantic coast are atypical of winter. The magnitude of the departure from winter condi-tions to surf conditions more typical of summer are de-monstrated by the occurrence of extreme frequencies for various surf directions during the period of the singularity (Table 1). Although Wahl[3] has shown that the pressure field characteristic of the singularity is present on January 20, the maxima and minima of frequencies for each surf direc-tion generally occur around January 23. This apparent lag may be attributable to the required time of wave generation and travel to the coast.

Table 1 Daily and 3-day extremes of surf directions along the US eastern coast during the January thaw

Station and surf direction	January thaw			Month of daily	
	Daily extreme (Date) (%)	3-day mean extreme (mid-date)	January daily mean (%)	maximum (% frequencies)	minimum
Point Judith, Rhode Island					
SE	38 (23rd)*	44 (22nd)*	55	May (65)	Sep (47)
SW	56 (22nd)§	43 (22nd)§	27	Aug (28)	May (19)
Atlantic City, New Jersey					
NE	0 (23rd)*	5 (23rd)	9	Dec (16)	Aug (1)
SE	57 (23rd)	63 (23rd)	77	Jul (84)	Mar (67)
S	19 (23rd)†	12 (23rd)†	3	May (7)	Oct (1)
Ocean City, Maryland					
NE	15 (21st)	18 (22nd)	23	Dec (30)	Jul (6)
E	28 (21st)	37 (22nd)	42	Dec (44)	Jul (26)
SE	52 (23rd)†	43 (22nd)†	30	Jul (58)	Dec (25)
Virginia Beach, Virginia					
NE	16 (23rd)*	26 (22nd)	39	Jan (39)	Jul (8)
SE	37 (23rd)†	28 (22nd)†	15	Jul (55)	Jan (15)
Oak Island, North Carolina					
SE	0 (22nd)	4 (21st)	5	May (13)	Jul (1)
S	74 (22nd)	69 (22nd)†	57	Oct (66)	Jun (52)
SW	26 (22nd)	25 (22nd)*	37	Apr (46)	Oct (24)
Hillsboro Inlet, Florida					
NE	0 (24th)*	5 (23rd)*	21	Jan (21)	Jul (1)
E	24 (23rd)‡	34 (23rd)	42	Apr (59)	Jan (42)
SE	47 (24th)§	42 (23rd)§	15	Jul (24)	Oct (7)

*Minimum value for month of January.
†Maximum value for month of January.
‡Minimum value for year.
§Maximum value for year.

At Ocean City, Virginia Beach, and Hillsboro Inlet, surf from the south-east dominates. At Atlantic City extreme high frequencies of south surf occur, and at Point Judith surf from the south-west dominates. On Oak Island southerly and south-westerly surf dominates, and at Hillsboro Inlet north-easterly and south-easterly surf predominates. The hypothesis enunciated here is not supported only at Atlantic City where the 1% level-of-significance criterion was not met.

Duquet[7] has suggested that the January thaw is a manifestation of "an adjustment of the planetary circulation from what might be called an early winter stage to a late winter stage . . . " The statistics of south-easterly surf at Hillsboro Inlet support the season-transition concept. The period of October 1–January 20 is characterised by a mean frequency of south-easterly surf of about 9%. During the period January 20–end March, the mean frequency of surf from the south-east is 23%. Thus, the time progression of south-easterly surf at Hillsboro Inlet takes the form of step functions rather than of a monotonic sequence.

Coastal wave climates are traditionally viewed as a simple annual cycle between summer and winter conditions, and definition within the annual cycle is not generally recognised. If other atmospheric singularities are found to mark simultaneous changes in wave climates, then such information would have immediate application to coastal planning in general, and coastal engineering in particular. Further investigation of coastal wave climates should answer this question and may further clarify the many unanswered questions about atmospheric singularities.

BRUCE P. HAYDEN

Department of Environmental Sciences,
University of Virginia,
Charlottesville, Virginia 22903

[1] *Glossary of Meteorology* (edit. by Huschke, R. E.), 512 (Am. met. Soc., Boston, 1959).
[2] Slocum, G., *Bull. Am. met. Soc.*, 22, 220 (1941).
[3] Wahl, E. W., *Bull. Am. met. Soc.*, 33, 380 (1952).
[4] Frederick, R. H., *Weather*, 21, 9 (1966).
[5] Brooks, C. E. P., *Weather*, 1, 107 (1946).
[6] Maejima, I., *Georg. Rep. Tokyo metrop. Univ.*, 2, 77 (1967).
[7] Duquet, R. T., *Mon. Weath. Rev.*, 91, 47 (Am. met. Soc., Boston, 1963).
[8] Wahl, E. W., *J. Met.*, 10, 4245 (1953).
[9] Lamb, H. H., *Geogrl J.*, 132, 183 (1966).

Middle Pleistocene stratigraphy in southern East Anglia

A STUDY of Quaternary deposits in south-east Suffolk revealed a complex, and hitherto unrecognised Middle Pleistocene stratigraphic sequence[1]. The probable ages of the units in the sequence are, from the base upwards, river sands and gravels of Beestonian age, a rubified palaeosol of Cromerian age, and a periglacial palaeosol, loess, glacifluvial gravel, and till of Anglian age (Table 1). A preliminary survey of the Quaternary deposits in Essex has revealed a similar sequence. We here outline the main characteristics of this sequence, which has now been recognised over some 2,000 km² in southern East Anglia, and discuss its palaeoenvironmental significance.

The Beestonian sands and gravels are the most widespread Pleistocene deposit in the region. They form a body which trends south-west–north-east across the area (Fig. 1) and reaches a maximum observed thickness of 30 m. They are characteristically though not invariably a whitish colour, and are composed predominantly of flint and quartzite with small percentages of other rocks such as chert, rhyolite and tuff[2]. Far travelled materials such as Bunter quartzite, chert and volcanic rocks show that much of the deposit has been transported into the region, and the absence of non-durable material such as chalk, London Clay mudballs, and Jurassic shells and shale from the main body of the deposit suggests that the bulk of the material has been subjected to prolonged fluvial abrasion. North-eastward dipping large scale and small scale cross-set structures with occasional silt lenses indicate that the material was deposited as bars and dunes by a river with an irregular regime and a north-eastward direction of flow. Intraformational ice-wedge casts, vein-ice casts, and segregated ice accommodation structures are developed throughout the deposit, demonstrating that the fluvial processes operated in a periglacial environment.

The height range of the surface of the deposit (Fig. 2) suggests that it is composite in origin, and may represent a series of river terraces formed at progressively lower levels as the river migrated south-eastwards. Considered in a regional context this body of sand and gravel probably
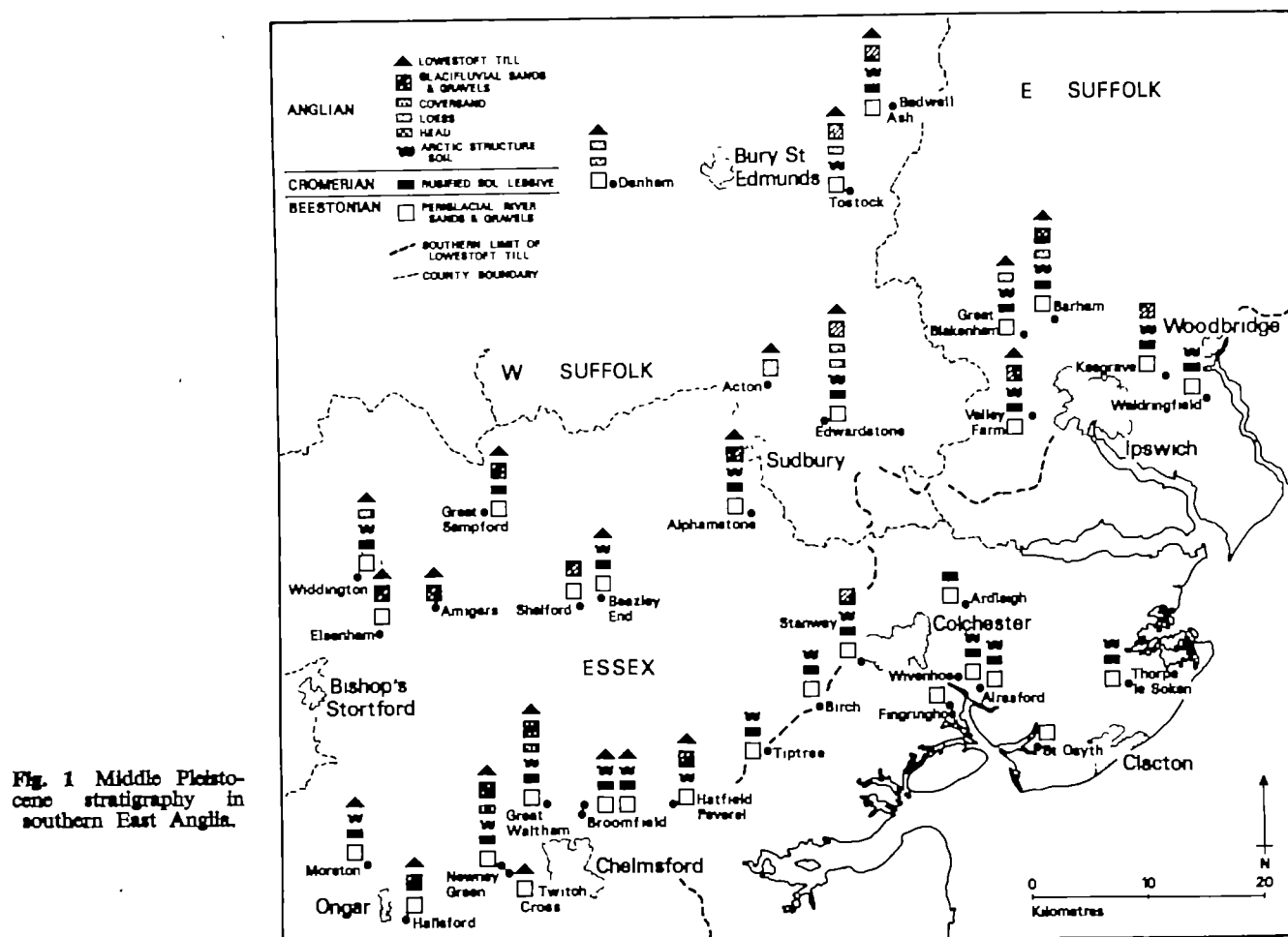
Fig. 1 Middle Pleistocene stratigraphy in southern East Anglia.

represents the debris of a much larger Thames drainage system, operating with a load and regime controlled by periglacial conditions.

These Beestonian sands and gravels were included by Prestwich[3] in his Westleton Beds which he considered to be marine in origin and separate from the glacial gravels. In contrast, Clayton[4] regarded all the sands and gravels of the Chelmsford and Harlow regions as glacial outwash, naming them the Danbury and Chelmsford Gravels, though Solomon[4] pointed out that the heavy mineral assemblage indicated that there were two separate units, one of which could not be associated with the local till. Bristow and Cox[5] also regarded these two gravels as glacial outwash but proposed that they both be called the Chelmsford Gravels.

The Cromerian palaeosol is a rubified *sol lessivé modaux*[6], and is developed on the surface of the Beestonian sands and gravels. Usually this soil is truncated, so that only the alluvial horizon survives in the form of a reddened zone associated with cutans, clay enrichment and iron enrichment. Soils of this type are produced in a humid warm

temperate climate[7,8], and as such reflect stable soil conditions which existed beneath the Cromerian mixed oak forest[9].

The Anglian palaeosol is an arctic structure soil[10], and is identified by involutions, frost cracks, and epigenetic ice-wedge casts which disturb the horizonation of the rubified *sol lessivé* and the primary structures in the upper part of the Beestonian sands and gravels. The loess that is developed in south-east Suffolk is represented by coversand in Essex (Fig. 1). The stratigraphic equivalence is demonstrated by a diagnostic heavy mineral assemblage (J. A. Catt, personal communication). Where both deposits exist at a site the coversand is uppermost. Their association with the arctic structure soil is shown by the fact that both deposits mix with the involutions and rest on the periglacial soil surface. At locations where the arctic structure soil has not been subsequently buried by the till, only the basal part of the ground ice structures remain, and the upper parts have been disturbed by younger periglacial activity.

The Anglian glacifluvial sands and gravels succeed the

Table 1 Middle Pleistocene stratigraphy and palaeoenvironments in southern East Anglia

Stage name	Formation name in south-east Suffolk	Environment	Sediments and palaeosols in Essex and west Suffolk
Anglian	Lowestoft Till	Glacial	Till (Lowestoft Till[11,15])
	Barham Sands and Gravels	Glacifluvial	Sands and Gravels
	Barham Loess Barham Arctic Structure Soil }	Periglacial	{ Loess and Coversand Head Arctic Structure Soil
Cromerian	Valley Farm Rubified *Sol Lessivé*	Humid, warm temperate	Rubified *sol lessivé*
Beestonian	Keagrave Sands and Gravels	Periglacial	Sands and Gravels (Essex White Ballast[8,11])

arctic structure soil. This deposit is only occasionally present beneath the till (Fig. 1), and beyond the till margin it is restricted to channels cut in underlying deposits. The base of the deposit is characterised by a pebble lag, but the main body consists of pale brown, well sorted sand and gravel composed predominantly of locally derived flint and quartzite (from the Beestonian sand and gravel) and a small component of glacially derived non-durable rocks. Primary sedimentary structures consist of small channels, plane beds, and small scale, tabular cross-sets which dip predominantly south-eastwards in Suffolk, and indicate that the deposit formed as dunes in shallow, braided rivers which drained away from the Lowestoft Till ice sheet in directions normal to the ice margin. Ground ice structures are absent from this deposit, and suggest that sedimentation rates were more rapid than the rate at which permafrost could develop. This deposit includes the upper part of the Chelmsford gravels[4].

Lowestoft Till of Anglian age[11,12] forms the upper part of the succession, and several of the sites have been studied by Perrin *et al.*[12]. At many locations the till rests directly on the undisturbed palaeosols or the wind-blown sediments. This indicates that over parts of southern East Anglia, glacial erosion did not occur, and probably suggests that flow tills protected the existing deposits before they were overridden by glacier ice.

The palaeoenvironmental significance of the Middle Pleistocene stratigraphic sequence in southern East Anglia can be summarised as follows:

(1) Two separate sand and gravel bodies can be recognised. The Beestonian river deposit is extensive throughout the region. The Anglian glacifluvial deposit is thin and localised. The earlier deposit is related to a north-eastward system of drainage. The present-day south-eastward and southward drainage patterns in the region were initiated during the Anglian glaciation with the deposition of the proglacial outwash.

(2) Downcutting and migration by the Thames drainage system began during the Beestonian periglacial stage, and by the time of the Cromerian Interglacial, the river must have adopted a route south and east of a line from Chelmsford to Colchester and Woodbridge (Fig. 2).

(3) The Cromerian rubified *sol lessivé* is developed widely across southern East Anglia. Where it is buried beneath Anglian glacigenic sediments it forms a clear stratigraphic horizon. Where it remains at the surface it forms the parent material for present-day soils which consequently have a reddened appearance.

(4) The Anglian arctic structure soil is developed across the region in association with loess and coversand. These aeolian deposits differ from Devensian aeolian sediments in their heavy mineral composition.

(5) Glacial erosion of pre-existing sediments and soils is negligible over large parts of the glacierised area of southern East Anglia.

J. ROSE

Department of Geography,
Birkbeck College,
University of London,
London W1A 1PA, UK

P. ALLEN

Department of Geology and Geography,
City of London Polytechnic,
Calcutta House,
Old Castle St,
London E1 7NT, UK

R. W. HEY

University of Cambridge,
Sedgwick Museum,
Downing Street,
Cambridge, UK

1 Rose, J., and Allen, P., *J. geol. Soc. Lond.* (in the press).
2 Hey, R. W., in *Quat. Res. Ass. Field Meeting Handbook, Clacton* (1973).
3 Prestwich, J., *Q. Jl. geol. Soc. Lond.*, **46**, 84–154 (1890).
4 Clayton, K. M., *Proc. Geologists Ass.*, **68**, 1–21 (1957).
5 Bristow, C. R., and Cox, F. C., *J. geol. Soc. Lond.*, **129**, 1–37 (1973).
6 Aubert, G., *Pedologie, Symp. intern.*, 3, 25–56 (1965).
7 Lamouroux, M., *Mem. O.R.S.T.O.M.* (Paris, 1972).
8 Birkeland, P. W., *Pedology, Weathering and Geomorphological Research* (Oxford University Press, New York, 1974).
9 West, R. G., in *The Vegetational History of the British Isles* (edit. by Walker, D., and West, R. G.) (Cambridge University Press, Cambridge, 1970).
10 Mückenhausen, E., *Enstehung, Eigenschaften und Systematik der Boden der Bundesrepublik Deutschland* (D.L.G. Verlag, Frankfurt, 1962).
11 Turner, C., in *Geol. Soc. Lond. Spec. Rep. No. 4* (edit. by Mitchell, G. F., *et al.*) (Geol. Soc. Lond., 1973).
12 Perrin, R. M. S., Davies, H., and Fysh, M. D., *Nature phys. Sci.*, **245**, 101–104 (1973).

Fig. 2 Elevations of surface of Beestonian sands and gravels. Locations are indicated on Fig. 1.
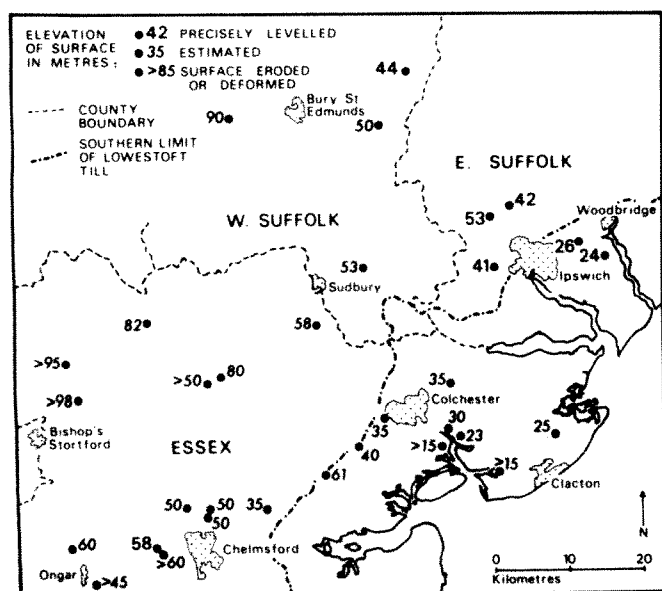
Oldest recorded *in situ* tracheids

RECENT critical reviews[1] have suggested that evidence from microfossils[2] (that is spores with triradiate marks, sheets of "cells" or tubes with tracheid-like thickenings) should not be considered sufficient to demonstrate the existence of vascular plants in Silurian times; the only acceptable

Fig. 1 Terminal globose sporangium (×18).
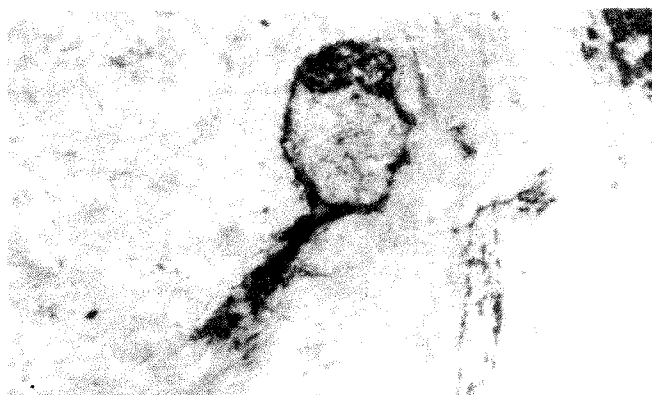
Fig. 2 Two spores with gaping trilete marks isolated on film pull and photographed using infrared film ($\times$750).

Fig. 4 Unbranched carbonaceous stem from which tracheids in Fig. 5 were isolated ($\times$7).

evidence being megafossils with tracheids *in situ*. Such was the evidence presented by Lang[3] when he described smooth axes from the Downton Series (Pridolian $\varepsilon\beta_2$) of the Welsh Borderland. These vascularised axes were found in association with two species of *Cooksonia*, very simple plants with smooth forking axes terminating in globose sporangia which contained spores bearing trilete marks. Megafossils morphologically similar to *Cooksonia* have been described from strata of similar age throughout the world, but the Lang specimens are the only ones with unequivocal tracheids and thus have hitherto been considered the earliest vascular plants. We report here the occurrence of a vascular plant in the slightly older Whitcliffian strata (Ludlow Series) of South Wales.

Our investigations have centred on a lower Whitcliffian flora (uppermost Ludlow Series$\equiv$Kopaninan $\varepsilon\beta_1$) from the Capel Horeb Quarry, Powys (SN 844323); the age of the plant-bearing Roman Camp Beds being based on faunal[4] and microfossil evidence (personal communication from K. Dorning). Among smooth often dichotomously branching carbonaceous axes are a few *Cooksonia*-type sporangia (Fig. 1) some of which contain spores with trilete marks (Fig. 2). A central longitudinal line visible in some of the axes is presumed to represent conducting tissue[5], although until the study reported here tracheids have not been isolated from this region. Regular transverse banding reminiscent of tracheidal pitting (Fig. 3) is often observed, but there is a possibility that such a pattern has been produced by the

regular fracturing of the carbonaceous residues. This explanation, however, cannot be applied to the short lengths of tracheids recovered on a cellulose acetate film pull from an unbranched axis 10 mm long and 1 mm wide (Fig. 4). Figure 5 shows two, possibly three adjacent incomplete tracheids, partly masked by a darker, more opaque, noncellular layer. The longitudinal walls are 7.5 μm apart in the longest element and the horizontal bars are, on average, 5 μm apart. Although we are reluctant to draw any conclusions on the type of pitting from such fragmentary compression fossils, there is some resemblance to the annular thickening described by Lang in his Downtonian specimens.
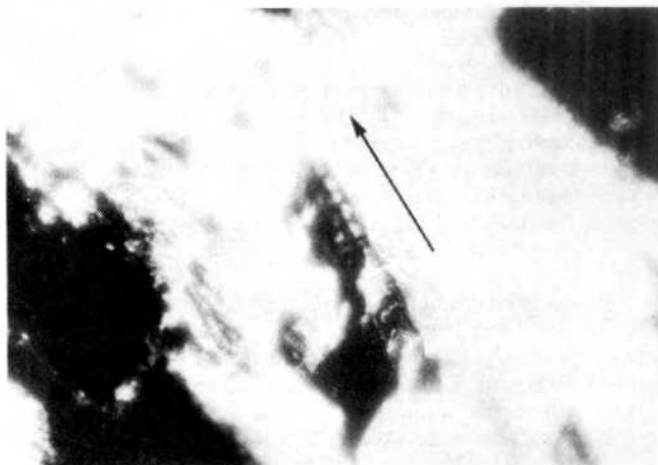
Fig. 5 Fragments of two adjacent tracheids isolated on a film pull and photographed using infrared film. Arrow indicates longitudinal axis of stem ($\times$416).

Our findings therefore indicate the presence of vascular plants, most probably of *Cooksonia*-type organisation, in the upper part of the Ludlow Series, slightly older than the Downtonian vascularised axes of Lang.

D. EDWARDS
E. C. W. DAVIES

Department of Botany,
University College,
PO Box 78,
Cardiff CF1 1XL, UK

1 Banks, H. P., *Rev. Palaeobot. Palynol.*, **20**, 13–25 (1975); *Bioscience*, **25**, 730–737 (1975).
2 Gray, J., Laufeld, S., and Boucot, A. J., *Science*, **185**, 260–263 (1974).
3 Lang, W. H., *Phil. Trans. R. Soc.*, **B227**, 245–291 (1937).
4 Potter, J. F., and Price, J. H., *Proc. geol. Ass., Lond.*, **76**, 379–402 (1965).
5 Edwards, D., *Palaeontology*, **13**, 3, 451–461 (1970).

Fig. 3 Possible tracheid showing regular transverse banding and a single longitudinal wall ($\times$416).

Wild bank voles (*Clethrionomys glariolus*) are possibly a natural reservoir of campylobacters (microaerophilic vibrios)

STRAINS of *Campylobacter*, a genus of microaerophilic bacteria, have been isolated by various workers from the gut and faeces of several animal species including man[1], but as far as we know, there have been no reports of their isolation from rodents. The continuing importance of diseases of farm animals involving campylobacters, for example infectious infertility and sporadic abortion in cattle[2], abortion in sheep[2] and dysentery in pigs[3], and the paucity of information on the source of outbreaks of some of these diseases prompted us to investigate whether campylobacters were present in the faeces of bank voles (*Clethrionomys glariolus*), short tailed field voles (*Microtus agrestis*) and long tailed field mice (*Apodemus sylvaticus*) from their natural habitat. These species are common throughout mainland Great Britain[4] and were being trapped as part of an investigation into the distribution of pathogenic bacteria and viruses in small wild rodents. Bank voles and field mice are commonly found in woodland and scrub, and short tailed field voles are found in grassland, therefore all three species might provide a source of infection for domestic animals. We report here that bank voles carry campylobacters.

At intervals of 6 weeks during a period of 3 months bank voles and field mice were trapped in an area of mixed broadleaved woodland in Berkshire, and short tailed field voles were trapped in an area of well established grassland in Hampshire. The animals were caught in Longworth traps[5] arranged in a squared grid of 3,600 m² in each trapping area. Two traps were placed at each point, with an interstation distance of 10 m. A faecal pellet voided by each animal was placed in 5 ml of quarter strength Ringer's solution in a 0.5-ounce McCartney bottle, and maintained at 18–20 °C for not more than 3 h before selective culture for campylobacters on blood agar by a filtration method[1]. Isolation plates were examined after incubation in 30% (v/v) CO_2 in air for 4 d at 37 °C. If colonies of *Campylobacter* were not detected, the plates were reincubated and examined after a further 10 d.

We isolated bacteria which we identified as *Campylobacter* from the faeces of 10 out of 13 bank voles, but not from the faeces of any of 12 field mice or any of 17 short tailed field voles examined. *Campylobacter* isolates from bank voles were motile, Gram-negative, slender curved rods with a predominance of coccal forms in old (5 d) cultures. On blood agar after incubation for 3 d at 37 °C colonies were irregular, tending to spread. Each colony had an entire edge and was light grey with a diameter of 2–4 mm. All isolates produced catalase and reduced nitrate to nitrite, but did not produce acid in media containing glucose, grow in nutrient gelatin or haemolyse blood. They produced H_2S detectable by lead acetate paper only when grown in a medium with added cysteine[6]. Growth did not occur, or was very slow, in media containing 1% (v/v) glycine and in media incubated either in air or in anaerobic conditions. By these criteria our isolates resemble *C. fetus* ssp. *venerealis* biotype *intermedius*[6], a resemblance which is interesting since campylobacters with these characteristics have been recognised as a cause of infectious infertility in cattle[7].

Our work shows that animals hitherto not considered in studies on the epidemiology of diseases of domestic animals carry campylobacters and may provide a reservoir of infection. Furthermore, it is possible that *Clethrionomys glariolus* might provide what is much needed for research into campylobacter infections—a laboratory model for pathogenicity studies.

D. S. FERNIE
T. D. HEALING

*Department of Microbiology,
University of Reading,
Reading RG1 5AQ, UK*

1 Morris, J. A., and Park, R. W. A., in *Isolation of Anaerobes* (edit. by Shapton, D. A., and Board, R. D.), 207–217 (Academic, London, 1971).
2 Lawson, J. R., in *Infectious Diseases of Animals: Diseases due to Bacteria* (edit. by Stableforth, A. W., and Galloway, I A.), 745–784 (Butterworth, London, 1959).
3 Fernie, D. S., Griffin, R. M., and Park, R. W. A., *Br. vet. J.*, 131, 335–338 (1975).
4 *Handbook of British Mammals* (edit. by Southern, H. N.), (Blackwell, Oxford, 1964).
5 Chitty, D., and Kempson, D. A., *Ecology*, 30, 536–542 (1949).
6 Véron, M., and Chatelain, R., *Int. J. syst. Bact.*, 23, 122–134 (1973).
7 Berg, R. L., Jutila, J. R., and Firehammer, B. D., *Am. J. vet. Res.*, 32, 11–22 (1971).

Population density affecting adult shell size of snail *Cepaea nemoralis* L.

THOMAS *et al.*[1,2] have questioned the ability of molluscs to limit their population size by self inhibition at high densities. They suggest that plant metabolites may have caused the reduction in growth and fecundity of the aquatic snail *Biomphalaria glabrata* Say ascribed to crowding in previous studies[3,4]. Field evidence for density-dependent regulation is limited. Yom-Tov[5] found that the fecundity of the desert snail *Trochoidea seetzeni* Pfeiffer was adversely affected by population density; he considered that either self-inhibition or nutritional differences were responsible. While studying the population dynamics and energetics of the land snail *Cepaea nemoralis* L., we have found evidence for density-dependent effects that cannot readily be explained by resource limitation. In samples of *C. nemoralis* collected in 1968 to study shell polymorphism (M. A. Palles-Clark, unpublished) a negative correlation was noticed between adult shell diameter and sample size ($r = -0.60$, $P < 0.001$). Snails were collected from chalk grassland between Beacon Hill and Round Down, West Sussex. A more detailed study carried out in the same area during the summer of 1973 confirmed this effect. The density of adult *C. nemoralis* was measured by mark-recapture at nine 20×20 m sites, situated 100–600 m apart. Each site was sampled in May, June and August. As there were insufficient recaptures to use multiple recapture methods for all sites, adult population densities were estimated for the second sampling occasion using the Lincoln index.

Measurement of adult shell diameter showed a 6–9% decrease in adult size associated with an increase in density from 0.5 to 5.5 adults per m² (Fig. 1). The pattern of population density was patchy but similar to that found in 1968. There were no simple relationships between density and geographical features, for example altitude or aspect. There were slight differences in shell colour and banding morph frequencies between the sites but these did not relate to differences in adult density or shell size.

Figure 1 also shows changes in mean shell diameter with season, reflecting a change in age structure of the adult population. Separate values of shell diameter are given for May, when the collections were mostly of old adults, surviving from previous years, and for August, when new adults (with shiny shells, periostracum intact) predominated. The August results exclude snails captured and marked in either May or June. Snails were considered adult when a lip had been formed at the mouth of the shell; after lip formation there is no further increase in shell diameter. Regression lines drawn through each set of data had significantly different slopes ($P < 0.02$), the change in slope indicating that recruitment to the low density populations increased mean adult size, whereas recruitment to the high density populations had the opposite effect. Long term
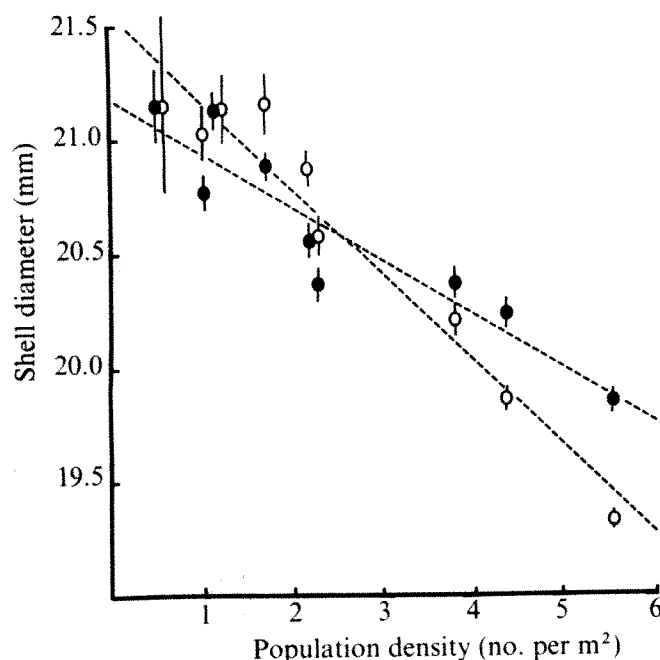
Fig. 1 Relationship of mean adult shell diameter ($\pm$s.e.) and adult population density. ●, Samples collected on May 23, $Y = 21.19 - 0.238X$ (s.e. slope $= 0.038$); ○, samples collected on August 7, $Y = 21.56 - 0.382X$ (s.e. slope $= 0.034$).

studies at an adjacent site did not show any significant changes in adult shell diameter over autumn and winter. It therefore seems that the differences between new and old adults are a function of the conditions experienced by juveniles, rather than selection pressures operating on adults. Unfortunately our field data do not enable us to separate the effects of juvenile and adult population densities on juvenile growth rates, nor to determine the time when growth is most affected.

The decrease in shell diameter with increasing density has a marked effect on body weight, reducing it by 14–31% (values determined from log–log regressions of weight and size for new and old adults[6]). Wolda's studies[7], in field conditions and in the laboratory, indicate that fecundity would be similarly reduced. His regressions relating shell diameter to clutch size and oviposition frequency for *C. nemoralis* predict a 22% decrease in annual egg production when adult shell diameter changes from 21.2 mm to 19.3 mm. Overwinter survival could also be affected: Cook and O'Donald[8] found that smaller adult *C. nemoralis* from a Pyrenean population were at a selective disadvantage when kept in experimental conditions producing high mortalities during dormancy.

A similar relationship between shell size and population density was detected by further analysis of Wolda's[9] field data. When adult shell diameter was plotted against density index the regression slope was negative and significantly different from zero ($P < 0.001$, combined data from east–west and north–south roads). Wolda's sampling sites were distributed over a total area similar to ours (about 0.5 km^2). Over larger areas, the variation in other factors affecting shell size, for example calcium availability, would be increased; consequently we would not expect the effect of population density to be so marked.

There was no indication that *C. nemoralis* at our study area was limited by either food quantity or quality. Point quadrat surveys of the nine sites did not show any correlation between population density and the vegetation biomass, nor between density and the proportion of palatable plants in the standing crop (R.A.D.C., P.W. and D. I. Morgan-Huws, unpublished). The feeding of *C. nemoralis* was studied in detail for a population with a

density of five adults per m^2 in June 1973, and a mean annual biomass (March 1973–March 1974) of 1.03 g m^{-2} for adults and juveniles. It was estimated[10] that this population consumed 3.7% of the above-ground production of herbs, 0.6% of that of grasses, and 0.91% of the total above-ground primary production. Preferred foods (dead herbs) were available at all times of the year. Calcium deficiencies were most unlikely: at all sites pieces of chalk were present at or near the soil surface. Shells of dead snails provided an alternative supply.

We therefore conclude that interactions between snails, either chemical or behavioural, were responsible for slowing juvenile growth rates in high density populations, resulting in smaller adults, and thus affecting the birth rate in succeeding years. It is likely that the changes in adult shell size that occurred in Wolda's field[9] and experimental[11] populations of *C. nemoralis* were caused by similar density-dependent effects. Laboratory studies are in progress to determine the effect of density on juvenile growth in experimental conditions. Initial results show that the presence of mucus trails of other snails inhibits activity, suggesting a possible mechanism for the effects observed in the field.

P. WILLIAMSON*
R. A. D. CAMERON†
M. A. CARTER

*Department of Biological Sciences,
Portsmouth Polytechnic,
Portsmouth PO1 2DY, UK*

1 Thomas, J. D., in *Advances in Parasitology*, 11 (edit. by Dawes, B.), 307–394 (Academic, London, 1973).
2 Thomas, J. D., Goldsworthy, G. J., and Aram, R. H., *J. anim. Ecol.*, 44, 1–27 (1975).
3 Berrie, A. D., and Visser, S. A., *Physiol. Zool.*, 36, 167–173 (1963).
4 Levy, M. G., Tunis, M., and Isseroff, H., *Nature*, 241, 65–66 (1973).
5 Yom-Tov, Y., *J. anim. Ecol.*, 41, 17–22 (1972).
6 Williamson, P., *J. anim. Ecol.* (in the press).
7 Wolda, H., *Arch. Neerl. Zool.*, 15, 381–471 (1963); *Evolution*, 21, 117–129 (1967).
8 Cook, L. M., and O'Donald, P., in *Ecological Genetics and Evolution* (edit. by Creed, R.), 93–108 (Blackwell, Oxford, 1971).
9 Wolda, H., *J. anim. Ecol.*, 38, 305–327 (1969).
10 Williamson, P., thesis, Portsmouth Polytechnic (1975).
11 Wolda, H., *Israel J. Malacol.*, 3, 63–71 (1972).

Overdominance and U-shaped gene frequency distributions

A GENERAL pattern that emerges from electrophoretic analyses of protein variation in natural populations is that alleles with intermediate frequencies are rare. This results, graphically, in a U-shaped frequency distribution when both allele frequencies from a two-allele locus are plotted. Supporters of the neutralist theory of population variation believe that most protein polymorphisms found in nature are the result of the random drift of neutral mutations and argue that a U-shaped distribution is expected for the case $4Ns < 1$, where N is the population size of diploid organisms and s is the average selection coefficient. We show here that, under certain assumptions, a U-shaped frequency distribution also obtains when $4Ns > 1$. From this we conclude that the analysis of allele frequency distributions (or, equivalently, relative contributions to heterozygosity) does not provide a sensitive means of resolving the neutralist and selectionist controversy.

In the textbook model of an overdominant polymorphism, two alleles, A and B, segregate at a locus, the frequency of allele A indicated by $\hat{p}$ and the genotypes AA, AB and BB have relative fitnesses $1 - s_1$, 1 and $1 - s_2$, respectively, where

$0 < s < 1$. It is easily shown that the equilibrium frequency of A, denoted $\hat{p}$, is $s_2/(s_1 + s_2)$. Various biochemical, physiological, genetic and ecological mechanisms can account for the relative superiority of the heterozygote. Our approach will specify the distribution of $\hat{p}$, given certain assumptions about the distributions of s_1 and s_2. We begin by postulating that s_1 and s_2 are independent, and will discuss the possibility of positive or negative correlation later.

The probability distribution of selection coefficients, $g(s)$, defined on the interval $0 < s < 1$, gives the probability that a selection coefficient will lie between s and $(s + ds)$. We assume a single parameter family of such distributions, where b is a parameter that reflects different biological reality (see below)

$$g(s) = \frac{1}{b} s^{[(1/b)-1]} \tag{1}$$

We now seek the probability distribution for equilibrium gene frequency, $h(\hat{p})$, given that both s_1 and s_2 are distributed according to equation (1). We define two new variables

$$x_1 = s_1/(s_1 + s_2) \tag{2a}$$
$$x_2 = s_1 + s_2 \tag{2b}$$

Solving for s_1 and s_2 gives

$$s_1 = x_1 x_2 \tag{3a}$$
$$s_2 = x_2(1 - x_1) \tag{3b}$$

The Jacobian of transformation (3), that is, the determinate of $\partial s_i / \partial x_j$, is x_2, thus the joint probability distribution of x_1 and x_2 is

$$H(x_1, x_2) = \left[\frac{1}{b}(x_1 x_2)^{[(1/b)-1]}\right]\left[\frac{1}{b}(x_2(1-x_1))^{[(1/b)-1]}\right]x_2 \tag{4}$$

Since x_1 and x_2 can be separated

$$H(x_1, x_2) = \frac{x_2^{[(2/b)-1]}}{b^2}(x_1(1-x_1))^{[(1/b)-1]} \tag{5}$$

It is then possible to obtain the marginal probability distribution of x_1, $h(x_1)$, by integrating over x_2 in the domain of the distribution of distribution function H ($0 \leq x_2 \leq 1, 0 \leq x_1 \leq \min(1/x_2, 1 - [1/x_2])$), yielding

$$h(x_1) = \begin{cases} x_1^{[(1/b)-1]}(1-x_1)^{[(-1/b)-1]}/(2b), & 0 \leq x_1 \leq 0.5 \tag{6a} \\ x_1^{[(-1/b)-1]}(1-x_1)^{[(1/b)-1]}/(2b), & 0.5 \leq x_1 \leq 1 \tag{6b} \end{cases}$$

This distribution, symmetrical about 0.5, gives the probability of $1 - \hat{p}$ ($= \hat{p}$) if s_1 and s_2 are taken from probability distribution equation (1). Functions $g(s)$ are plotted in Fig. 1A for different values of b, with the corresponding $h(\hat{p})$ distributions plotted in Fig. 1B. The question remaining is which, if any, values of b correspond to biological reality.

Although a fairly high percentage of newly arisen mutations may be lethal, most of the alleles persisting at a locus in a natural population are not seriously deleterious as homozygotes: really deleterious genes will be eliminated soon after they appear. Thus it is reasonable to postulate that homozygotes that persist in a population will have fitness nearly equal to the heterozygote and that the mode of the distribution $g(s)$ will be at $s = 0$, with the probability of larger s falling off rapidly with increasing s. This corresponds to large values of b, which give U-shaped gene frequency distributions.

As mentioned, we have assumed that s values are uncorrelated. A negative correlation, which pairs large and small s values,
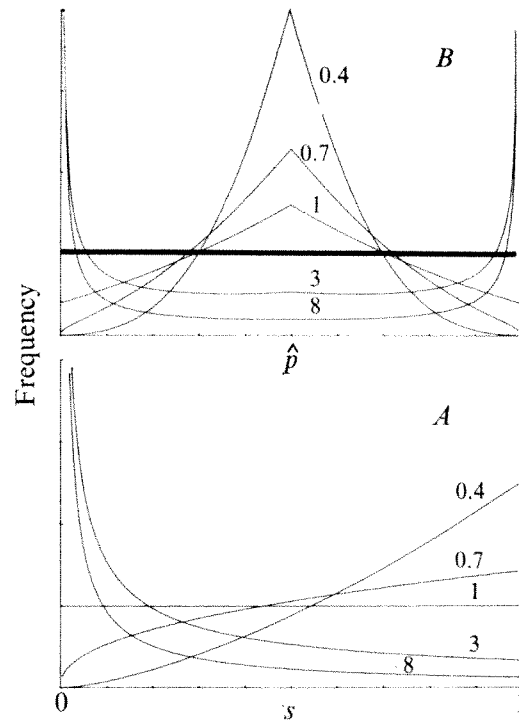
Fig. 1 *A*, Some $g(s)$ distributions for different values of the parameter b, the numerical value of which is given above the line. *B*, Some $h(\hat{p})$ distributions for the same five values of parameter b. The heavy line is a hypothetical distribution for which all values of $\hat{p}$ are equally likely.

seems to us absurd; it would, however, increase the tendency towards U-shaped gene frequency distributions. Positive correlations between s values work against U-shaped distributions. A likely source of such a positive correlation is a situation where the selection coefficients for the homozygotes at some 'important' loci tend to be high, whereas those for other 'less important' loci tend to be low. We admit that some loci are more essential than others, but alleles persisting in natural populations at such relatively essential loci cannot deviate far from optimal functioning before the s values become large. Conversely, populations seem to tolerate a much broader range of function at loci which *a priori* are less critical metabolically. For example, an $s = 0.5$ might correspond to a very slight biochemical change in a regulatory enzyme on the one hand, and to a virtual absence of function of a coat colour gene on the other. We feel therefore, that it is incorrect to equate physiological or developmental importance with the effect of a locus on fitness. Rather, it is reasonable that the s distributions for all types of loci are the same and that our assumption that the individual s values at a locus are uncorrelated is warranted.

This result is not solely dependent on equation (1). We have used other single parameter probability distributions for $g(s)$. One family of functions was

$$g(s) = b(1 - s^b)/(b + 1) \tag{7}$$

We were not able to obtain $h(p)$ analytically for this function, so we used numerical integration on a computer. Distributions of $g(s)$ which were concave towards the origin (corresponding to values of b in equation (7) less than 1) gave U-shaped distributions for the equilibrium gene frequency. In addition, we treated other $g(s)$ distributions in a Monte Carlo fashion and arrived at the same conclusion. If the natural distribution of selection coefficients is concave towards the origin, then U-shaped probability distributions of gene frequencies are to be expected.

Using equation (1) we have performed Monte Carlo simulation for more than two alleles at a locus. For tri-allele systems,

where the gene frequencies of the population can be represented as the three distances from the sides to an interior point in an equilateral triangle, the probability distribution for the equilibrium surface can be plotted as a surface above the triangular domain. Values of b that give U-shaped gene frequency distributions for two-allele cases give three-allele gene frequency distributions that have modes in the three corners and very low probabilities in the centre of the triangular domain. The same is true for four or more alleles.

The foregoing results apply to infinite populations. The situation for finite populations, where some sampling error (drift) is unavoidable, can be sketched. Robertson[1] introduced the idea of a retardation factor, which shows the strength of an overdominant polymorphism for slowing fixation. This depends on both the size of a population and on the deterministic frequency of equilibrium. For population sizes greater than 1,000, equilibrium frequencies in the range 0.2–0.8 are quite secure—the mean time to fixation is very long. When the equilibrium frequency is near 0 or 1, however, the rate of fixation may be even faster than for a purely neutral case.

By tending to fix those equilibrium frequencies in the ranges 0.05–0.20 and 0.80–0.95, such drift would change the A-shaped distributions in Fig. 1B to W-shaped distributions. It would also tend to square off the corners of a U-shaped distribution, but it would be impossible statistically to distinguish the two cases. Therefore, if our arguments are correct, gene frequency data cannot be used to determine whether most electrophoretic polymorphisms are neutral or selected.

Yamazaki and Maruyama[2] concluded that gene frequency data, in the form of contribution to relative heterozygosity, was inconsistent with the overdominant model. But this conclusion was based on the assumption that all equilibrium frequencies of overdominant polymorphisms are equally probable (corresponding to the heavy horizontal line in Fig. 1B); whereas, following the reasoning outlined in this paper, we believe that this assumption is not reasonable and that their conclusion is unwarranted.

We have not explained the phenomenon of the U-shaped allele frequency distribution. We have merely shown that it may be consistent with a selectionist model.

MICHAEL GILPIN
MICHAEL SOULÉ
ANATOLA ONDRICEK

Department of Biology

ELIZABETH A. GILPIN

School of Medicine,
University of California, San Diego,
La Jolla, California 92093

1 Robertson, A., *Genetics*, **47**, 1291–1300 (1962).
2 Yamazaki, T., and Maruyama, T., *Science*, **178**, 56–58 (1972).

Triploid pseudogamous biotype of the leafhopper *Muellerianella fairmairei*

THE sibling leafhopper species *Muellerianella fairmairei* (Perris, 1857) and *M. brevipennis* (Boheman, 1847) (Homoptera, Delphacidae) are widely sympatric in Europe. Studies of the biological differences between the two species, and of their hybridisation has revealed a remarkably high proportion of females of *M. fairmairei* sampled in the field and a gradual disappearance of males in laboratory colonies. *M. brevipennis* had a normal sex ratio in field samples and when reared in the laboratory[1]. Additional cytogenetic studies now demonstrate that in *M. fairmairei* a pseudogamous female triploid biotype coexists with the diploid bisexual form.

Populations of *M. fairmairei* (Southern Greece and Southern France) and *M. brevipennis* (Holland), which maintain a 1:1 sex ratio are diploid ($2n = 28$). In both species the male

karyotype consists of thirteen autosome pairs and one XY pair ($2n = 13$ AII + XY). The female karyotype (Fig. 1a and b) shows thirteen autosome pairs and a large pair of XX chromosomes ($2n = 13$ AII + XX).

Populations of *M. fairmairei* from Holland (Leersum) and England (Ascot–Silwood Park) with a high proportion (75–100%) of females seemed to comprise diploid males and a mixture of diploid and triploid females. Somatic metaphases of the triploid females, as far as could be determined, consistently had 41 chromosomes, including three X chromosomes. Ovarian semi-mature and mature eggs always contained a metaphase plate with 41 "pseudobivalents" (Fig. 1c). Their size was approximately half that of the bivalents of the diploid female, and they resembled secondary oocyte chromosomes. In this insect the first and second meiotic divisions probably do not take place and are replaced by a single equational division.

After the triploid biotype of *M. fairmairei* had been isolated in the laboratory, three experimental crosses were made to investigate the role of the sperm in the production of triploid females. When the cross was *M. fairmairei* (3n) ♀♀ × *M. fairmairei* ♂♂, five pairs reared on their host plant, *Holcus lanatus* (L.) produced triploid females only and in high numbers (160–514) females per pair. All five females appeared to have mobile sperm in their spermathecae. When the cross was *M. fairmairei* (3n)♀♀ × *M. brevipennis* ♂♂, six pairs placed in a cage containing *H. lanatus* and *Deschampsia caespitosa* (L.) (the latter is the specific food plant of *M. brevipennis*) produced comparatively few triploid females (a total of 88). Three of the six crossed females had mobile sperm in their spermathecae. When the cross was *M. fairmairei* (3n)♀♀ × sterile hybrid ♂♂, nine sterile males (produced by previous crossing of *M. brevipennis* (2n)♀♀ × *M. fairmairei* ♂♂) were paired with females in glass tubes containing *H. lanatus*. Although the triploid females

Fig. 1 The chromosome complements of *M. fairmairei* (2n): a, *M. brevipennis*; b, *M. fairmairei* (3n); c, at the first female meiotic division, which occurs in the nucleus of semi-mature eggs. The bar in (c) represents 10 µm. All photomicrographs are reproduced at the same magnification.
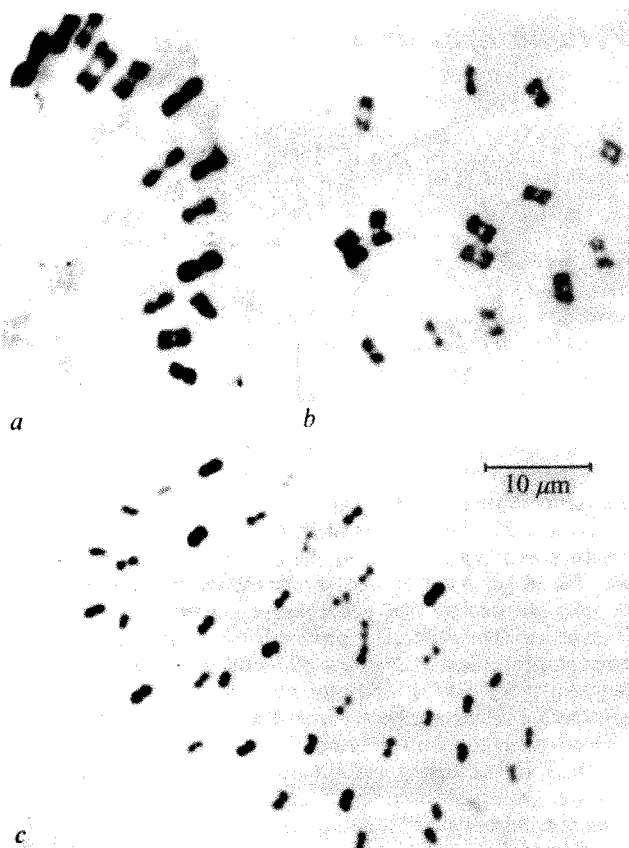
were observed several times copulating with the hybrid males, none of the 163 eggs deposited exhibited embryogenesis.

Thus the triploid biotype of *M. fairmairei*, which is morphologically indistinguishable from the diploid female, has a typically pseudogamous mode of reproduction. The entrance of sperm (derived from the males of *M. fairmairei* or the closely related *M. brevipennis*) is necessary for embryonic development to occur, but the sperm probably do not fuse with the egg nucleus. Copulation with sterile males is ineffective. So far this mode of reproduction has not been reported in Hemiptera[2], although other forms of parthenogenesis are well known in this order.

When both female biotypes of *M. fairmairei* were reared together in the presence of males, the diploid type disappeared after one or two generations. The higher fecundity of the pseudogamous biotype and the fact that it produces all-female progeny possibly explain this phenomenon. It is not known how both biotypes can coexist in the field. It is clear, however, that the triploid type requires the availability of a diploid population for its maintenance.

An indication that the triploid biotype might be of hybrid origin is its ability to oviposit and develop on *D. caespitosa*, the host plant of *M. brevipennis*. Diploid hybrids, originated from crosses of *M. fairmairei* $(2n)$♀♀ $\times$ *M. brevipennis* ♂♂ and the reciprocal cross, have the same ability. Also, the meiotic behaviour of some female diploid hybrids (28 "pseudobivalents") closely resembles that of the triploid.

I thank Professor J. van der Veen and Drs J. Sybenga, R. H. Cobben and J. Gut for discussions.

SAKIS DROSOPOULOS

Laboratory of Entomology,
Agricultural University,
Wageningen, The Netherlands

1 Drosopoulos, S., *Ent. Berichten*, 35, 154–157 (1975).
2 White, M. J. D., *Animal Cytology and Evolution*, 961 (Cambridge University Press, Cambridge, 1973).

Somatic hybridisation of *Petunia hybrida* and *P. parodii*

INTRASPECIFIC and interspecific fusion of plant protoplasts is induced by agents[1-4]. Protoplasts of several species can regenerate a cell wall and divide to form callus, from which whole plants can be grown[5]. To combine these processes and obtain hybrid cells and then hybrid plants several requirements must be met[6]. The parental protoplasts should be capable of being cultured, as should the hybrid cells, and plant regeneration must be possible. Selection procedures developed so far[7-10] may not be generally applicable, and we have now developed a method which makes use of naturally occurring differences in the sensitivity of cultured plant protoplasts to growth media and drugs[11].

One half of our selection procedure was based on a difference in growth between leaf protoplasts of *Petunia hybrida* and *P. parodii*. Protoplasts of *P. parodii*[12] never grow beyond the small colony stage (about 50 cells) in a medium similar to that of Murashige and Skoog[13] (M/S medium) while protoplasts of *P. hybrida* cultivar Comanche[14] produced callus[15]. The other half of the selection procedure was based on a difference in sensitivity to actinomycin D, *P. hybrida* protoplasts being more sensitive than those of *P. parodii*. The selection procedure is shown in Fig. 1.

Leaf protoplasts of *P. parodii* and *P. hybrida* were suspended in 9% (w/v) mannitol solution containing inorganic salts[17] (2×10^5 protoplasts per ml) and dispensed into screw-capped tubes in 8-ml volumes so that there was a viability control, and a fusion viability control for each species. A mixture of equal volumes of each species was used for the fusion treatment.

All tubes, except viability controls, were centrifuged (80*g*,

10 min) and the supernatant was removed. To each tube was added 2 ml of 15% (w/v) polyethylene glycol (Koch-Light, molecular weight 6,000), 4% (w/v) sucrose and 0.01 M CaCl₂ to induce fusion. Protoplasts were resuspended, left for 10 min at 25 °C and then gradually diluted by the addition, at 5-min intervals, of the M/S medium in volumes of 0.5 ml, 1.0 ml, 2.0 ml, 2.0 ml, 3.0 ml and then 4.0 ml. Protoplasts were resuspended after each addition. After centrifugation at 60*g* for 15 min, the supernatant was replaced by M/S medium (8 ml per tube). At this stage the two viability controls were centrifuged and the protoplasts were resuspended in M/S medium. A further control was prepared by mixing equal volumes of *P. parodii* and *P. hybrida* fusion viability controls (post-fusion mixture control). All tubes were left for 1 h before plating. On average 4% of the protoplast population had fused as determined by counts of nuclei.

Protoplasts were maintained in liquid medium by the liquid-on-agar culture method[15]. To each 9-cm plastic Petri dish was added 8 ml of M/S medium containing actinomycin D (1.0 µg ml⁻¹), solidified with 0.5% (w/v) agar. To the surface of the agar was added 4 ml of M/S medium, containing actinomycin D (2.0 µg ml⁻¹), plus 4 ml of the appropriate protoplast suspension (at 2.0×10^5 protoplasts per ml), in M/S meduim without actinomycin D so that the final concentration of the drug in the liquid layer was 1.0 µg ml⁻¹, with a protoplast density of 1×10^5 per ml. Protoplasts of the two viability controls were plated in M/S medium without actinomycin D, as were samples of the *P. parodii* and *P. hybrida* fusion controls. Dishes were maintained at 27 °C with continuous illumination of 1,000 lx from daylight fluorescent tubes.

After 28 d, colonies were transferred to dishes containing M/S medium with 3% (w/v) mannitol solidified with 1% (w/v) agar and no actinomycin D. Growing colonies were transferred finally after 60 d to M/S medium containing no mannitol. Colonies were also recovered from the *P. hybrida* fusion control, maintained in the absence of actinomycin D, and from the *P. parodii/P. hybrida* fusion treatment in the presence of actinomycin D. All other dishes were devoid of colonies except the *P. hybrida* viability control. After a further 10 weeks, the small calluses of the *P. parodii/P. hybrida* fusion treatment, and the *P. hybrida* fusion control, were transferred to M/S medium, containing IAA (2.0 µg ml⁻¹) and 6-benzylaminopurine (1.0 µg ml⁻¹), for plant regeneration. Ten calluses of separate origin were recovered from the *P. parodii/P. hybrida* fusion treatment. Of these, eight produced shoots on this medium. Subsequently they were removed and rooted in M/S medium containing NAA (0.1 µg ml⁻¹) as the sole growth regulator and 0.3% agar. Plantlets thus produced were potted and grown up to flowering. Callus from the *P. hybrida* fusion control did not regenerate shoots on the same medium, although some shoots were produced when zeatin (1.0 µg ml⁻¹) was added.

No plants of *P. parodii* were recovered since the protoplasts did not grow beyond the colony stage. Plants produced from the *P. hybrida* fusion control were all diploid ($2n = 14$), red flowered and like the parent species. All plants produced from *P. parodii/P. hybrida* fusion were purple flowered, and thus distinguishable from the parents (Fig. 2). Chromosome counts on roots showed a minimum number of 24, and a maximum of 28 in these plants.

Seedlings of *P. parodii* and *P. hybrida* were treated with colchicine (0.3% (w/v) aqueous solution) for 90 min to produce tetraploid ($4n = 28$) plants. The F₁ sexual hybrid (*P. hybrida* $\times$ *P. parodii*) was obtained at the tetraploid level using colchicine. The reciprocal cross (*P. parodii* $\times$ *P. hybrida*) was identical in all respects. As Fig. 2 shows, flowers of the plants with 28 chromosomes, produced by fusion, were very similar to those of the tetraploid F₁ sexual hybrid. Tetraploid *P. parodii* and *P. hybrida* parents were

Petunia parodii

Petunia hybrida

Leaf Protoplasts → Leaf Protoplasts

M/S Medium

M/S medium + actinomycin D
(1.0 μg ml⁻¹)

Cell colonies

No further growth

M/S medium

Callus

M/S medium + actinomycin D
(1.0 μg ml⁻¹)

No growth

Fusion

M/S medium
+ actinomycin D (1.0 μg ml⁻¹)

Somatic hybrid callus

Somatic hybrid plants
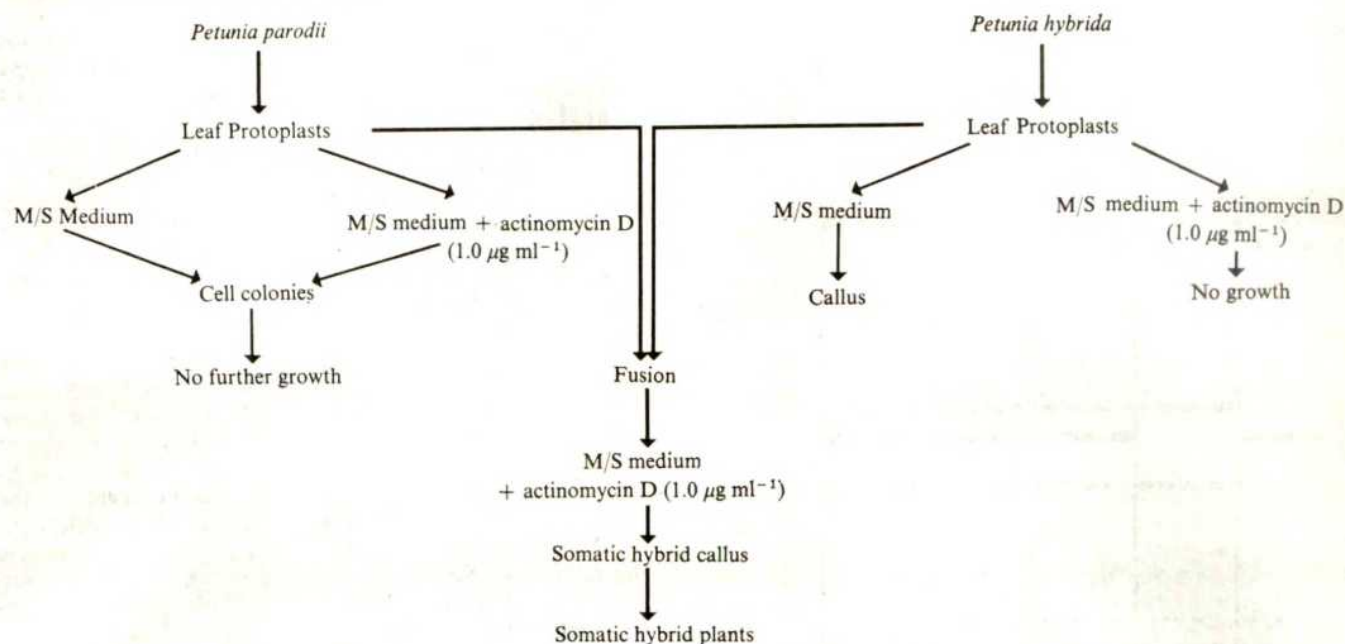
Fig. 1 Selection scheme for the production of somatic hybrids of *P. parodii* and *P. hybrida* based on differential parental protoplast growth responses to M/S medium and actinomycin D. Both species gave high yields of protoplasts (2–3×10^6 per g fresh weight) with plating efficiencies of up to 40% in M/S medium, with naphthalene acetic acid (NAA) (2.0 μg ml⁻¹), 6-benzylaminopurine (6-BAP) (0.5 μg ml⁻¹), sucrose (3%) (w/v) and mannitol (9%) (w/v), pH 5.8. Division of *P. hybrida* protoplasts was limited by actinomycin D at 0.75–1.0 μg ml⁻¹ in M/S medium, yet protoplasts of *P. parodii* were not affected up to 5.0 μg ml⁻¹. In animal cells, actinomycin D resistance is dominant[16], and likely also to be so in *Petunia*. Callus of *P. parodii*[12] readily undergoes shoot formation, yet that of *P. hybrida* does not[15]. It was expected that the regeneration potential of somatic hybrids would be provided by the *P. parodii* genome. In M/S medium plus actinomycin D protoplasts of *P. hybrida* would be expected to remain alive, but not divide while protoplasts of *P. parodii* would divide to the small colony stage. If complementation occurs in the somatic hybrid, the *P. hybrida* genome would permit division beyond the colony stage, while the *P. parodii* genome would confer resistance to actinomycin D. Protoplasts did not cross feed with respect to the M/S medium and actinomycin D growth responses.

easily distinguished from the F_1 sexual hybrid and the regenerated plants, for doubling of the chromosomes in the parents affected neither flower colour nor morphology. Therefore, the plants resulting from fusion that had 28 chromosomes were classed as somatic hybrids (*P. hybrida* + *P. parodii*). The sexual hybrid ($4n$) and the somatic hybrid (28 chromosomes) were both self-fertile and cross-fertile.

Determination of leaf isoperoxidase showed that the somatic hybrids had the spectrum of bands associated with both parents (diploid and tetraploid); and 20% of stained gels had an extra faint band found also in the tetraploid F_1 sexual hybrid (Fig. 3). The presence of this additional band, coupled with the hybrid nature of the flowers excludes the possibility that a chimaera had been obtained.

Fraction 1 proteins of *P. hybrida* and *P. parodii* were purified and analysed by A. A. Gatenby in our laboratory, using a specific immunoabsorbent, as described before[18]. Preliminary isoelectric focusing experiments of S-carboxy-methylated protein, in the presence of 8 M urea, have revealed an identical polypeptide composition of the chloroplast-coded large subunit for both species. The small nucleus-coded subunit of *P. hybrida* contains two polypeptides[19]. *P. parodii* also contains two nucleus-coded polypeptides, both with isoelectric points identical to those found in *P. hybrida*. This is perhaps expected, for within the genus *Petunia* there seems to be a high degree of chromosome homology[20,21].

Experiments based on the same selection principle, and culture procedure, but with different concentrations of actinomycin D, have given selected calluses. Somatic hybrid plants have been produced after selection in M/S medium

Fig. 2 Flowers of (from top to bottom) tetraploid *P. parodii* (white), tetraploid F_1 hybrid (*P. hybrida* × *P. parodii*) (purple), somatic hybrid (28 chromosomes) (purple) and tetraploid *P. hybrida* (red). The purple flower colour corresponds to *Fuschia* purple (BCC No. 199, 28/1) and the red corresponds to Turkey red (Horticultural Colour Chart, 721). The somatic hybrid was also distinguishable from either of the parents on the basis of corolla tube and peduncle length. Petals of tetraploid *P. hybrida* were fringed, yet those of the F_1 hybrid and the somatic hybrid were smooth. Abaxial vein pigmentation of the floral bud was pronounced only in the F_1 hybrid and the somatic hybrid. Pollen of both parents was yellow, whereas that of the F_1 and somatic hybrids was occasionally purple.

Fig. 3 Diagram of starch gel electrophoresis of peroxidase isoenzymes from leaf material of: *a*, *P. hybrida* (4*n*); *b*, *P. parodii* (4*n*); *c*, *P. hybrida* (4*n*) and *P. parodii* (4*n*) mixed before homogenisation; *d*, *P. hybrida* × *P. parodii* (4*n*); *e*, somatic hybrid (*P. hybrida* + *P. parodii*) (28 chromosomes). Leaf material was homogenised in 0.2 M sodium chloride solution at 4 °C using 1 g fresh weight of leaf per 0.5 ml of solution. A sample of the fresh homogenate was absorbed into a filter paper wick which was inserted into a freshly made 10% (w/v) starch horizontal slab gel (Electrostarch). The gel buffer contained 90% (v/v) 0.01 M citric acid and 0.065 M Trizma base, *p*H 8.2, plus 10% (v/v) 0.025 M lithium hydroxide and 0.2 M boric acid *p*H 7.9. The tank contained only the latter buffer. After electrophoresis for 6 h at 4 °C (40 mA), the gel was sliced horizontally, and a lower slice was stained for peroxidases using 100 mg of *O*-dianisidine in 70 ml of 95% (v/v) ethanol, plus 28 ml of acetate buffer and 2 ml of 3% (w/v) hydrogen peroxide solution. (The acetate buffer contained 0.14 M sodium acetate and 0.06 M acetic acid.) Staining was conducted at room temperature for 45–60 min, followed by washing in 70% (v/v) ethanol; photographed, and R_f values recorded. ★, Position of additional band.

with actinomycin D at 0.9 µg ml⁻¹, 0.8 µg ml⁻¹ and 0.75 µg ml⁻¹. At 0.75 µg ml⁻¹ the selection conditions allow some *P. hybrida* callus to grow through. In these circumstances, somatic hybrids were preferentially recovered since the shoot-inducing medium only rarely supports shoot formation in *P. hybrida*. So far, four separate experiments have produced somatic hybrid plants. Calluses have been selected in each of two further experiments. These calluses have regenerated shoots but have not yet been analysed fully. In one experiment which produced no somatic hybrids, actinomycin D (2.0 µg ml⁻¹) was added 10 d after fusion treatment, and *P. hybrida* was not sensitive to this concentration at the colony stage, and in the presence of *P. parodii*.

The frequency of somatic hybrid plant formation was variable but was of the order of one in 1–8 × 10⁵ protoplasts. This does not take into account loss of protoplast viability

on plating or losses due to the effect of the fusion treament.

Because neither mutant nor haploid plants are required, this selection strategy should be readily applicable to a wide range of species. We are investigating—using this complementation-selection procedure based on naturally occurring differences between species—the consequences of fusion of protoplasts isolated from sexually incompatible species.

J. B. Power
E. M. Frearson
C. Hayward
D. George
P. K. Evans
S. F. Berry
E. C. Cocking

ARC Group,
Department of Botany,
University of Nottingham,
University Park,
Nottingham NG7 2RD, UK

1 Power, J. B., Cummins, S. E., and Cocking, E. C., *Nature*, **225**, 1016–1018 (1970).
2 Keller, W. A., and Melchers, G., *Z. Naturforsch.*, **28C**, 737–741 (1973).
3 Kao, K. N., and Michayluk, M. R., *Planta*, **115**, 355–367 (1974).
4 Wallin, A., Glimelius, K., and Eriksson, T., *Z. Pflanzenphysiol.*, **74**, 64–80 (1974).
5 Nickell, L. G., and Heinz, D. J., in *Genes, Enzymes and Populations* (edit. by Srb, A. M.), 109–128 (Plenum, New York, 1973).
6 Power, J. B., and Cocking, E. C., in *Applied and Fundamental Aspects of Plant Cell, Tissue and Organ Culture* (edit. by Reinert, J., and Bajaj, Y. P. S.), (Springer, Berlin, in the press).
7 Carlson, P. S., Smith, H. H., and Dearing, R. D., *Proc. natn. Acad. Sci. U.S.A.*, **69**, 2292–2294 (1972).
8 Melchers, G., and Labib, G., *Molec. gen. Genet.*, **135**, 277–294 (1974).
9 Melchers, G., Keller, W., and Labib, G., in *Modification of the Information Content of Plant Cells* (edit. by Markham, R., Davies, D. R., Hopwood, D. A., and Horne, R. W.), 161–168 (North-Holland, Amsterdam, 1975).
10 Gleba, Y. Y., Butenko, R. G., and Sytnik, K. M., *Dokl. Akad. Nauk. S.S.S.R.*, **221**, 1196–1198 (1975).
11 Cocking, E. C., *et al.*, *Pl. Sci. Lett.*, **3**, 341–350 (1974).
12 Hayward, C., and Power, J. B., *Pl. Sci. Lett.*, **4**, 407–410 (1975).
13 Murashige, T., and Skoog, F., *Physiol. Pl.*, **15**, 473–497 (1962).
14 Frearson, E. M., Power, J. B., and Cocking, E. C., *Devl Biol.*, **33**, 130–137 (1973).
15 Power, J. B., *et al.*, *Pl. Sci. Lett.*, **7**, 51–55 (1976).
16 Sobel, J. S., Albrecht, A. M., Riehm, H., and Biedler, J. L., *Cancer Res.*, **31**, 297–307 (1971).
17 Cocking, E. C., and Peberdy, J. F. (eds), *The Use of Protoplasts from Fungi and Higher Plants as Genetic Systems*, 67 (Department of Botany, University of Nottingham, 1974).
18 Gray, J. C., and Wildman, S. G., *Pl. Sci. Lett.*, **6**, 91–96 (1976).
19 Chen, K., Kung, S. D., Gray, J. C., and Wildman, S. G., *Pl. Sci. Lett.* (in the press).
20 Mather, K., *J. Genet.*, **45**, 215–235 (1943).
21 Sink, K. C., *Can. J. Genet. Cytol.*, **27**, 67–74 (1975).

Pea root nodules containing more than one *Rhizobium* species

PARTICULAR species of legume are nodulated only by appropriate species of *Rhizobium*. Thus *R. leguminosarum*, *R. phaseoli* and *R. trifolii* nodulate peas, French beans and clover, respectively. The basis of this specificity between the two partners of the symbiosis is not understood. Complete specificity is not always found, however. Although *R. trifolii* does not normally nodulate peas, induced mutants of this species able to nodulate this host have been described[1] and some strains of *Rhizobium* isolated from clover nodules nodulated both clover and peas[2]. Selected clover lines have been nodulated by *R. leguminosarum*[3]. It has been reported that a non-rhizobial contaminant can occur in soybean nodules[4] and that the presence of *R. trifolii* may allow *R. meliloti* to occupy clover root nodules (B. O. Gillberg, personal communication). We show here that the presence of *R. leguminosarum* can allow *R. trifolii* and *R. phaseoli* to enter pea nodules but that *R. trifolii* cannot fix nitrogen in the nodules of this 'foreign' host.

The strains of *R. trifolii* (6001 and 6100) and *R. phaseoli* (1233) we used did not normally nodulate peas in our conditions. Of 15 plants inoculated with strain 1233 and 43 with strain 6001 none had any nodules, while one out of 36

plants inoculated with R. trifolii strain 6100 had three very small ineffective nodules. Since no viable bacteria could be isolated from them after surface sterilisation, strain 6100 could not be identified; conceivably they arose by contamination. The R. trifolii and R. phaseoli strains effectively nodulated clover and French bean plants, respectively.

The results of inoculating peas with various strains of R. leguminosarum together with R. trifolii or R. phaseoli are shown in Table 1. For each mixed inoculum at least one nodule on at least half of the plants contained the heterologous species. The frequency of such nodules was 2–27%. All nodules occupied by R. trifolii or R. phaseoli also contained the appropriate strain of R. leguminosarum. In most mixed nodules the heterologous species represented the minority of the bacteria recovered, but in some there were many more R. trifolii than R. leguminosarum. For example, from a nodule on a plant inoculated with strains 6100 plus 6099, 2.8×10^8 R trifolii and 2.2×10^2 R. leguminosarum were recovered.

For all mixed inoculations of R. trifolii plus ineffective strains of R. leguminosarum there was no nitrogen fixation as determined by the absence of acetylene reduction whether R. trifolii was the majority or minority strain. Four of the R. leguminosarum strains (6009, 6099, 7111 and 7112) were derived from single mutagenic treatments from effective parents and thus can be assumed to be defective in one or very few steps. It is apparent that the defect(s) in these strains and of the two ineffective field isolates of R. leguminosarum (922 and 983) could not be overcome by the R. trifolii strains Strain 922 dies not inhibit acetylene reduction when it occurs in the same pea nodule as an effective strain of R. leguminosarum[1].

A limited number of white clover (Trifolium pratense) plants was inoculated with R. trifolii strain 6001 plus R. leguminosarum strain 922 (which did not nodulate clover on its own). Of six nodules obtained from two plants, one contained both 922 and 6001. All six nodules were pink and effective, confirming that the presence of strain 922 in nodules did not inhibit nitrogen fixation by effective strains. This also demonstrates that occupation of nodules by heterologous species of Rhizobium is not restricted to peas.

Two independent R. leguminosarum strains (6007 and 6008) unable to nodulate peas were derived after ultraviolet

treatment of an effective strain (897). To determine whether the presence of non-infective strains of R leguminosarum would enable R. trifolii to occupy pea nodules, peas were inoculated with R. trifolii (strain 6001) plus strains 6007 or 6008; no nodules were formed. Strains 6007 or 6008 do not inhibit nodulation by effective strains of R. leguminosarum. Neither of these strains was recovered from any of 237 nodules from plants inoculated with nodulating strains of R. leguminosarum plus strains 6007 or 6008. It is perhaps surprising that the presence of an infective strain of R. leguminosarum does not allow non-infective mutants of R. leguminosarum to enter pea nodules but does permit entry of strains of R. trifolii or R. phaseoli.

It should be noted that the techniques used here did not enable us to determine whether the heterologous species formed bacteroids in pea nodules since only the undifferentiated bacteria can form colonies on plates. Nor was it possible to determine where the heterologous strain occurred within the nodules. Transfer of the plasmid RP4 between strains of R. leguminosarum has been observed in pea nodules[1,2], but it was not found in any of the thirteen nodules that contained both R. trifolii and the RP4-carrying strain of R leguminosarum (6099). RP4, however, is transferred by conjugation in vitro from strains of R. leguminosarum to strains 6001 and 6100 at much lower frequencies than it is to strains of R. leguminosarum (A.W.B.J., unpublished observations). The failure to find transfer in the nodules therefore need not imply that the two strains were physically separated.

It has been reported that the ability to nodulate clover can be transferred by conjugation from R. trifolii to R. phaseoli[1] and by transformation from R trifolii to R. meliloti[2]. In this study the results were not explicable by transfer of the ability to nodulate peas from R leguminosarum to strain 6001 or 6100 since single colony isolates of these strains obtained from pea nodules retained their ability to nodulate clover but did not nodulate peas.

The bases of the specificity in the recognition between legumes and Rhizobium may reside in a barrier or barriers that are breached only by bacteria of the appropriate cross-inoculation group. By definition, infective strains of R. leguminosarum overcome these barriers in peas. A general explanation for the results presented is that the hetero-

---

Table 1 Incidence of pea nodules containing R. phaseoli (strain 1233) or R. trifolii (strains 6001 and 6100) in the presence of strains of R. leguminosarum

Heterologous strain	R. leguminosarum strain		Effective (E) or ineffective (I)	No. of plants examined	No. of plants with heterologous strain in one or more nodules	No. of nodules examined	No. of nodules with heterologous strain
	Stock no.	Genotype					
1233	7482	phe str-r	E	2	1	22	3
6001	897	phe trp str-r	E	2	2	49	3
6001	6009	phe trp str-r	I	2	2	29	2
6001	7111	phe	I	2	1	22	1
6100	7111	phe	I	2	1	12	1
6001	7112	phe	I	1	1	12	2
6001	6099	spc-r (RP4)	I	1	1	43	2
6100	6099		I	5	5	58	11
6001	922	his str-r	I	2	1	45	1
6100	922		I	3	3	33	9
6001	983	str-r	I	2	1	30	1
6100	983		I	2	1	20	2

R. phaseoli strain 1233 is a rifampicin-resistant mutant of the pigment (probably melanin)-producing strain 3644 in the Rothamsted collection. R. trifolii strains 6001 (rif-r) and 6100 (rif-r str-r) were derived from strain CC37 supplied by E. A. Schwinghamer. R. leguminosarum strain 922 (his rif-r) was derived from ineffective strain 330 obtained from B. Gillberg. Strain 983 (rif-r) was derived from the ineffective Rothamsted strain 1012. All other R. leguminosarum strains were derived from our wild-type effective strain 300 (ref 6): 6099 (spc-r) carries the plasmid RP4 and is ineffective; 603 (phe) was the parent of 897 (phe trp str-r) and 7482 (phe str-r); 6009 is an ineffective mutant obtained after ultraviolet treatment of 897; 7111 and 7112 were independent ineffective mutants obtained after nitrosoguanidine mutagenesis of strain 603. Culture and inoculation methods have been described before[6]. Peas (var. Dark Skin Perfection) were inoculated with about 2×10^8 of the inoculant strains. After about 4 weeks, nodules were excised, surface sterilised and crushed in 20% glycerol. Samples (0.04 ml) of the macerates were plated on suitable selective media[6,9]. The macerates were stored frozen and those that contained the heterologous strain were diluted and plated on selective media to allow enumeration of the two strains.

logous *Rhizobium* species use the apparatus induced by *R. leguminosarum* so that there is a partial breakdown in specificity. The strains of *R. trifolii* and *R. phaseoli* used do not ordinarily induce nodules on peas but we do not know if they form infections which abort before nodule formation. *R. trifolii* was unable to fix nitrogen in pea nodules. This suggests that more than one barrier prevents *Rhizobium* of one cross-inoculation group forming an effective symbiosis with a legume of a different group. This is in accord with the findings that mutants of *R. trifolii* able to nodulate peas did not fix nitrogen in this host[1] and that some strains of *Rhizobium* isolated from clover formed ineffective nodules on peas[2].

ANDREW W. B. JOHNSTON
JOHN E. BERINGER

*Department of Genetics,
John Innes Institute,
Colney Lane, Norwich NR4 7UH, UK*

1 Schwinghamer, E. A., *Am. J. Bot.*, 49, 269–277 (1962).
2 Kleczkowska, J., Nutman, B. P. S., and Bond, G., *J. Bact.*, 48, 673–675 (1944).
3 Hepper, C. M., *Rothamsted Experimental Station Report*, 282 (1975).
4 Jansen van Rensburg, H. J., and Strijdom, B. W., *Phytophylactica*, 3, 125–130 (1971); *ibid.*, 4, 1–8 (1972).
5 Johnston, A. W. B., and Beringer. J. E., *J. appl. Bact.*, 40, 375–380, (1976).
6 Johnston, A. W. B., and Beringer. J. E., *J. gen. Microbiol.*, 87, 343–350 (1975).
7 Higashi, S., *J. gen. appl. Microbiol.*, 13, 391–403 (1967).
8 Laine, R. T., and Alexander, M., *Can. J. Microbiol.*, 7, 959–961 (1961).
9 Beringer, J. E., *J. gen. Microbiol.*, 84, 188–198 (1974).

Erucic acid, an accidental additive in bread

RAPESEED oil is a stable, viscous oil utilised extensively by the baking industry in several countries as a lubricant to prevent the dough piece sticking to the container during baking. The oil is applied either to the container or directly on to the dough surface and allows the easy removal of the bread after baking. There are, however, few data available to indicate the extent to which any rapeseed oil is absorbed into the bread during processing. The question has assumed considerable importance since Roine[1] and many subsequent workers[2-4] showed that a range of laboratory animals developed heart lesions and other biochemical abnormalities when fed diets with a substantial content of rapeseed oil. These effects have generally been attributed to erucic acid[5] (13-docosenoic acid) which is the major fatty acid constituent in normal rapeseed oils (30–50% of total fatty acids)[6]. Consequently, various countries have legislated to restrict the amount of erucic acid permitted in certain foods, although no tolerances have been set specifically for baked products. This is mainly due to the assumption that erucic acid is not a natural component of the ingredients used in bread.

We have examined 20 different types of bread obtained from retail outlets in Sydney, Australia and found that all loaves contained erucic acid in concentration which ranged from 170 to 1,070 mg of erucic acid per kg of fresh loaf, with an average concentration of 405 mg kg^{-1}. A sample of wheat flour was also analysed but no erucic acid was detected (<0.2 mg kg^{-1}), and so the erucic acid in the breads can be assumed to be derived from the rapeseed oil applied just before baking. The breads analysed included plain and fancy rolls, and standard loaves of white, brown, wholemeal, rye, kibble wheat, soy flour–wheat and starch-reduced breads. The breads were dried and ground, and a sample (20 g) was extracted with solvent to remove the lipids which were saponified and methylated[7], and analysed by flame ionisation gas chromatography using a 10% DEGS–PS on Supelcoport (80/100) glass column at 190 °C with a carrier gas flow rate of 20 ml per min of nitrogen.

Measurements were also made of the distribution of erucic acid throughout a single loaf. Table 1 shows that most of the erucic acid was present in the outer layer

Table 1 Distribution of erucic acid in a bread loaf (15 cm high)

Position in loaf	mg erucic acid per kg
Base–0.2 cm into loaf	9,550 ⎫
	⎬ 3,320*
0.2–1.0 cm	1,760 ⎭
1.0–3.0 cm	50
3.0–7.5 cm (centre of loaf)	40

*Overall value for bread from base to 1.0 cm.

although some was found in the centre of the loaf. If these gradients can be applied to all loaves, then the crust slice (1 cm thick) would contain appreciable amounts of erucic acid and range from about 0.15 to 1.0% (w/w) of the crust.

Although much information has been obtained during the past 5 yr on the effects of erucic acid on animals, negligible data are available on its toxicological effects on man, and so it has been difficult to set meaningful levels of erucic acid permissible in human diets. An acceptable daily intake of rapeseed oil has been estimated to be 30 mg per kg of body weight[8], which is equivalent to about 12 mg of erucic acid per kg. Our data show that an adult would achieve this intake if of the order of 1 kg of bread was eaten per day, although more than half the daily allowance would be supplied by two crust slices. Children, however, could receive more than their total allowable intake with only two crust slices. Thus, the use of rapeseed oil as a release agent in the baking industry leads to contamination of the product with erucic acid of a magnitude sufficient to cause concern for public health. The continued use of this substance on bread should therefore be reviewed. There are rapeseed oils readily available on the international market with low erucic acid contents ($<1\%$)[9] and these could be investigated for use in the baking industry. It is probable, however, that the viscosity of these oils would not be as desirable as that of high-erucic rapeseed oils, in which case, blending with other components would be necessary.

R. B. H. WILLS
M. WOOTTON
G. HOPKIRK

*School of Food Technology,
University of New South Wales,
Kensington, Australia 2033*

1 Roine, P., Uksila, E., Teir, H., and Rapola, J., *Z. ErnahrWiss.*, 1, 118–124 (1960).
2 Rocquelin, G., and Cluzan, R., *Ann. Biol. anim. Biochim. Biophys.*, 8, 395–406 (1968).
3 Abdellatif, A. M. M., and Vles, R. O., *Nutr. Metab.*, 12, 296–305 (1970).
4 Walker, B. L., Lall, S. P., Slinger, S. J., and Bayley, H. S., *Proc. int. Conf. Sci. Tech. Marketing Rapeseed Prod.*, Ste Adele, Canada, 377–404 (1970).
5 Abdellatif, A. M. M., and Vles, R. O., *Nutr. Metab.*, 15, 219–231 (1973).
6 Hilditch, T. P., and Williams, P. N., *The Chemical Constitution of Natural Fats*, fourth ed., 294–303 (Chapman and Hall, London, 1964).
7 Oulaghan, S. A., and Wills, R. B. H., *N.Z. Jl exp. Agric.*, 2, 381–383 (1974).
8 *Inf. Bull. Br. ind. biol. Res. Ass.*, 11, 272–273 (1972).
9 Downey, R. K., *Chem. Ind.*, 401–406 (1976).

Suppression of adenocarcinoma by the immunological consequences of calorie restriction

EARLIER extensive studies have indicated that calorie restriction as well as protein and amino acid restriction inhibit the spontaneous development of mammary or lung adenocarcinomas, hepatomas and certain chemical carcinogen-induced tumours in rodents[1-4]. Jose and Good[5-7] and Cooper *et al.*[8] have shown that chronic moderate protein deprivation in mice and rats dramatically depresses antibody production while increasing, or permitting maintenance of, vigorous cell-mediated immune responses. The latter include abilities to resist virus infection, reject skin allografts and develop killer-cell activity

Table 1 Rate of appearance of mammary tumours (MT) and mortality in C3H female mice given diet with normal or restricted calories

| Calories per day | No. of mice dead with tumours per total dead Days | | | | | Total dead MT/Total | % |
	101–200	201–300	301–400	401–500	501–600		
16	0	0	4/5	7/11	1/1	12/17	71
10	0/4	0	0/4	0/4	0/6	0/18	0

against syngeneic and allogeneic tumour cells. More profound chronic protein deprivation, however, depresses both cell-mediated and humoral immunity[7,9]. Walford et al.[10,11] have shown that calorie restriction delays development of immune functions at an early age but prolongs maintenance of immunological capacity and survival in long lived, tumour-free mice. Further, Fernandes et al.[12] have shown that dramatic prolongation of life of autoimmunity-prone, short lived mice was produced by life-long calorie restriction, and that tumours did not appear in these mice. Although some of the findings clearly link moderate protein and calorie restriction to heightened cellular immunity, no definitive efforts have yet been made to compare directly the influence of dietary restriction on spontaneous tumour development and immunological functions. This report presents observations on the influence of calorie restriction on development of spontaneous mammary adeno-

the development of spontaneous mammary adenocarcinoma in the female C3H/Umc mice, and more than 50% of them lived longer than 400 d. In contrast, 71% of the control group of female C3H/Umc mice on the normal (16 calories) diet developed mammary tumours by 500 d.

In Table 2 a comparison of mitogen response studies carried out by using the technique described previously[13] is presented for the spleen cells obtained from mice on the two diets. Responsiveness of spleen cells to optimum concentrations of the T-cell lectins, phytohaemagglutinin (HA17) or concanavalin A (con A) was not only maintained on the low calorie intake but actually increased. The responses to the B-cell mitogen lipopolysaccharide (LPS) were also maintained at a level at least equivalent to that achieved by spleen cells of mice in the control group.

Table 3 summarises data comparing capacity to develop

Table 2 The incorporation of ^{3}H-TdR with and without mitogenic stimulation of spleen cells from normal and calorie restricted C3H/Umc 3-month-old mice

| Calories per day | Cultures* (h) | Controls | c.p.m. of ^{3}H-TdR (at optimum concentration of mitogens†) | | |
			+PHA (2.5 µg)	+con A (.5 µg)	+LPS (10 µg)
16	48	770±240	100,164±11,327	76,834±16,825	42,018±7,740
	72	467±170	74,369±4,597	101,274±20,959	20,051±4,669
10	48	1,636±194	144,018±11,087	103,752±10,354	43,122±734
	72	671±226	80,712±5,732	154,559±6,218	21,360±1,836

*0.5 × 10⁶ spleen cells from 4 individual mice per group cultured in RPMI-1640 media (triplicates) with foetal calf serum (2%) for 48 or 72 h at 37 °C in CO_2 and humidity in 3,040 microtest II tissue culture plates. Cultures were labelled with tritiated thymidine ^{3}H-TdR (0.5 µCi per well) for an additional 16 h.
†Source of mitogens—PHA (HA17), Wellcome; con-A, Sigma, LPS, Difco.

carcinoma in mice and relates the suppression of tumour development observed to alterations in immune functions produced by calorie restriction at weaning.

Four-week-old female C3H/Umc mice were fed diets comprised of 22% casein, 33% dextrose, 33% starch, 5% corn oil, 4% salt mixture and 2% vitamin mixture, as reported previously[5,8,13]. The protein-calorie-deprived mice were fed 10 calories instead of the 16 calories eaten by the control mice, but in all other respects the diets were the same. The food was provided in conditions in which each mouse consumed its entire ration each day. The influence of this form of chronic calorie restriction on survival of C3H/Umc mice is shown in Table 1. In mice of this strain, calorie restriction did not significantly alter lifespan, even though calorie-restricted mice lived longer (3/18) than any of those in the control group. Table 1 shows that the diet low in calories completely prevented

direct haemolytic plaque-forming cells (PFC) 4 d following (intraperitoneal) immunisation with sheep red blood cells (SRBCs)[14]. After the primary stimulation, the number of direct PFCs developed in the spleens of mice on a low calorie diet was markedly depressed as compared with that of the mice in the normal calorie group. In contrast, after a secondary stimulation, the PFC formation was not significantly different in the two groups. Of interest was the finding that when the spleen cells of both groups were compared for ability to generate a secondary PFC response after injection into lethally X-irradiated (950 R) recipients[15], the sensitised cells from the calorie-deprived donors showed a greater number of PFC per million cells than was generated by spleen cells from sensitised mice which had been fed the higher calorie diet. Although several interpretations for this observation are possible, a likely explanation is that a suppressor cell population active in the

Table 3 Influence of normal and caloric restriction on immune response to SRBCs in C3H/Umc mice

| Calories per day | Primary response* (direct PFCs) | | Secondary response† (indirect PFCs) | | PFC response in X-irradiated (950 R) host‡ | |
	10⁶	Spleen	10⁶	Spleen	10⁶	Spleen
16	127±17 (7)	10,131±1,123	231±121 (5)	19,625±9,223	95±22 (10)	2,058±749
10	40±15 (6)	2,091±984	181±13 (5)	13,327±2,409	190±34 (10)	1,668±404
Student's t test	P < 0.01	P < 0.01	NS	NS	P < 0.02	NS

*Immunised (intraperitoneal) with 10 × 10⁶ SRBCs for 4 d.
†Primed mice for 7 d with SRBC were again immunised and PFC assay was carried out on day 4.
‡Spleen cells (30 × 10⁶ or 40 × 10⁶) from immunised (primary) normal or low calorie mice were injected along with SRBCs into the lethally X-irradiated 8-week-old mice fed a normal laboratory diet. PFC assay was done on day 4 or 5 after injection.
Figures in parentheses represent the number of mice tested in each group.

Table 4 Antigen induced inhibition of DNA (^{125}IUdR)* synthesis in normal and calorie-restricted C3H 3-month-old mice

Calories per day	Body weight (g)	Injected with	Spleen weight (mg)	Per 100 mg spleen (c.p.m. ±s.e.)	% change
16	31±2 (6)	Saline	54±2	2,115±13	
16	32±2	SRBCs	79±7	2,952±406	+39
10	22±1 (6)	Saline	48±3	3,112±36	
10	23±1 (6)	SRBCs	68±3	1,817±174	−42

*2 µCi of ^{125}IUdR (Amersham/Searle, specific activity 106 µCi µg^{-1}, Batch No. 294BA) was injected intraperitoneally 6 h after administration (intravenously) of saline or SRBCs. Twenty hours later mice were killed, spleens weighed and isotope incorporation was determined in a Gamma Counter (Nuclear Chicago).

calorie-deprived host is not as effective as it was in the intact animal at inhibiting immune response to SRBCs after transfer to the lethally X-irradiated syngeneic recipients. The irradiated mice given sensitised spleen cells from the calorie-deprived mice showed no increase, however, in absolute number of PFCs per spleen (Table 3). This might be attributed to differences in relative proportions of different subpopulations of donor cells and their capacity to migrate or to cooperate during the immune response in the spleen of X-irradiated host animals.

The technique of Zatz and Goldstein[16], which may be one means to evaluate a population of suppressor T cells by their capacity to suppress DNA synthesis *in vivo* after injection of SRBCs, was then used to determine whether the suppressor T-cell influence is indeed excessive in the spleens of the calorie-deprived C3H/Umc mice. Table 4 shows that, by this analysis increased suppressor effect for this T-cell dependent response was also demonstrable in the calorie-deprived animals. It can be seen from Table 4 that, whereas no suppressor-cell influence was demonstrable by this technique in the spleens of 3-month-old mice fed the normal calories, those fed the low calorie diet showed very marked evidence of T-cell suppressor activity.

Our findings confirm those of several investigators[1-4] in showing that calorie restriction *per se* prevents development of mammary adenocarcinoma in highly susceptible C3H/Umc mice. They also show that moderate protein and calorie restriction permits maintenance of a vigorous functional thymic cell population, as has been found with chronic protein deprivation[5,6,8,13]. Calorie restriction apparently also decreases the PFC response to T-dependent antigen in 3-month-old mice. That such a marked deficiency of PFC response to SRBCs is not due to the changes in the functional capacity of macrophages is suggested by data from a previous study[8] which indicated that macrophage function is normal in protein–calorie deprived mice. Perhaps this issue still needs to be studied more extensively. The basis of the humoral immunodeficiency in the primary response of the calorie-deprived mice might be associated *in vivo* with an endocrine hypofunction[17] which may permit a T-cell suppressive influence in the host. This interpretation is suggested by the observation that spleen-cell response *in vitro* to both PHA, con A and LPS was not altered by calorie deprivation, and that secondary PFCs in response to SRBCs are produced quite well when the previously stimulated spleen cells of the calorie-deprived animals are transferred into lethally irradiated hosts. This postulated relationship is compatible with the recent evidence presented with "T-dependent" and "T-independent" antigens, suggesting that a T-cell population present either in adult or newborn spleens possesses the ability to recognise feedback signals, and makes a decision to suppress or enhance the immune response on the basis of the amounts and molecular form of antigen given either in a suppressive or non-suppressive environment[18,19]. More direct evidence for an active suppressor-cell population in the calorie-deprived mice, as has been revealed in young NZB mice[20], was then obtained by an assay which measures DNA synthesis in the spleen. The association of this immunological perturbation with resistance

to mammary tumours and with the marked change in hormonal functions[21] and their influence on target tissues or receptor sites of a host given a calorie-restricted diet deserves extensive study.

An interesting possibility suggested but not proved by these findings is that T-suppressor cells might suppress directly the development of mammary cancer. This is a remote possibility but it should be studied directly. Another possibility is that the calorie-deprived mice whose lymphoid system is so much perturbed might lack the ability to develop still another form of suppressor cells. This latter cell type has recently been shown to occur in animals during tumour growth, and is implicated in the inhibition of cellular immune functions[22-25]. The elimination of the development of such suppressor-cell activity and/or absence of serum-blocking factors[5,6] by calorie restriction might facilitate the destruction of newly arising tumour cells by immunocompetent T lymphocytes and macrophages. The nature and mode of such immunological alterations in relationship to dietary restriction remain unknown, and further studies are urgently required to elucidate the role of nutritional manipulations in susceptibility and resistance to tumour development *in vivo*. Similarly, although these studies and previous investigations from our laboratories seem to suggest that immunity could be involved in the resistance to tumours produced by low calorie diet, much further study will be needed to elucidate these relationships.

GABRIEL FERNANDES
EDMOND J. YUNIS

Department of Laboratory Medicine and Pathology,
University of Minnesota,
Minneapolis, Minnesota 55455

ROBERT A. GOOD

Sloane–Kettering Cancer Institute,
New York, New York 10021

1 Rous, P., *J. exp. Med.*, **20**, 433–451 (1914).
2 Visscher, M. B., Ball, Z. B., Barnes, R. H., and Sivertsen, I., *Surgery*, **11**, 48–55 (1942).
3 Tannenbaum, A., in *The Physiopathology of Cancer* (edit. by Homburger, F.), 517–526 (Hoeber-Harber, New York, 1959).
4 Carroll, K. K., *Cancer Res.*, **35**, 3374–3383 (1975).
5 Jose, D. G., and Good, R. A., *Nature*, **231**, 323–325 (1971).
6 Jose, D. G., and Good, R. A., *Cancer Res.*, **33**, 807–812 (1973).
7 Jose, D. G., Stutman, O., and Good, R. A., *Nature*, **241**, 57–58 (1973).
8 Cooper, W. C., Good, R. A., and Mariani, T., *Am. J. clin. Nutr.*, **27**, 647–664 (1974).
9 Jose, D. G., and Good, R. A., *J. exp. Med.*, **137**, 1–9 (1973).
10 Walford, R. L., Liu, R. K., Gerbase-Delima, M., Mathies, M., and Smith, G. S., *Mech. Ageing Dev.*, **2**, 447–454 (1973–74).
11 Gerbase-Delima, M., Liu, R. K., Chenoy, K. E., Mickey, R., and Walford, R. L., *Gerontologia* (in the press).
12 Fernandes, G., Yunis, E. J., and Good, R. A., *Proc. natn Acad. Sci. U.S.A.*, **73**, 1279–1283 (1976).
13 Fernandes, G., Yunis, E. J., and Good, R. A., *J. Immun.*, **116**, 782–790 (1976).
14 Bullock, W. W., and Möller, E., *Eur. J. Immun.*, **2**, 514–517 (1972).
15 Makinodan, T., Sado, T., Groves, D. L., and Price, G., *Curr. Top. Microbiol. Immun.*, **49**, 80–113 (1966).
16 Zatz, M. M., and Goldstein, A. L., *J. Immun.*, **110**, 1312–1317 (1973).
17 Huseby, R. A., Ball, Z. B., and Visscher, M. B., *Cancer Res.*, **5**, 40–46 (1945).
18 Gershon, R. K., Orbach-Arbouys, S., and Calkins, C., *Prog. Immun.*, **2**, 123–133 (1974).

[19] Mosier, D. E., and Johnson, B. M., *J. exp. Med.*, **141**, 216–226 (1975).
[20] Dauphinee, M. J., and Talal, N., *J. Immun.*, **114**, 1713–1716 (1975).
[21] Zumoff, B., Fishman, J., Bradlow, H. L., and Hellman, L., *Cancer Res.*, **35**, 3365–3373 (1975).
[22] Gershon, R. K., Mokyr, M. B., and Mitchell, M. S., *Nature*, **250**, 594–596 (1974).
[23] Gorczynski, R. M., *J. Immun.*, **112**, 1826–1838 (1974).
[24] Kuperman, O., Fortner, G. W., and Lucas, Z. J., *J. Immun.*, **115**, 1282–1287 (1975).
[25] Fujimoto, S., Greene, M. I., and Sehon, A. H., *J. Immun.*, **116**, 800–806 (1976).

Mechanisms of bone destruction in the development of skeletal metastases

SKELETAL metastases are associated with progressive bone destruction but the pathogenesis is not well understood. I report here experiments which indicate that two principal mechanisms are involved. The first is mediated by osteoclasts and the tumour secretes a diffusable osteoclast activating factor; the second functions when the residual trabeculae are surrounded by the tumour cells.

Several hundred specimens of bone, obtained at autopsy from the spines of 68 patients who died from various malignant diseases[1], were fixed in neutral formalin and decalcified in 5% nitric acid until no residual calcification was seen radiographically. Sections were then stained with haematoxylin and eosin and examined microscopically.

A suspension of VX2 carcinoma was obtained by finely dividing approximately 1 g of tumour in 5 ml of 0.15 M NaCl, producing approximately 2.5×10^8 cells per ml. In the first group a 0.5-ml sample of the suspension was injected into one tibia or ilium of New Zealand white rabbits or on to the periosteal surface. With tibial implants (16 rabbits), the bone was exposed at operation under ether anaesthesia, and a fine drill hole, just large enough to accommodate a size 18 needle, was made through one cortex. The tumour cell suspension was injected well down the medullary cavity using a bent size 18 needle. In the first few experiments the hole was plugged with bonewax, but this was found to be unnecessary and was discontinued. Iliac implantation was made by percutaneous injection into the iliac crest (six rabbits) or into a hole drilled in the ilium at open operation (six animals); all the periosteal injections were percutaneous (12 rabbits).

In half the animals 0.15 M NaCl was injected into the contralateral bone; in the others no operation was performed on this bone. The rabbits were killed at intervals from 24 h to 8 weeks later, and specimens from both the injected and the contralateral bone were treated in the same way as the human necropsy specimens. In the second group of 33 rabbits, 1 ml of tumour cell suspension was used by the other experimental details were the same.

Fig. 1 Metastasis from prostatic carcinoma. There is considerable new bone formation (B) in the stroma of the tumour. Multiple osteoclasts (arrowed) can be seen, adjacent to areas of bone destruction ($\times 96$).
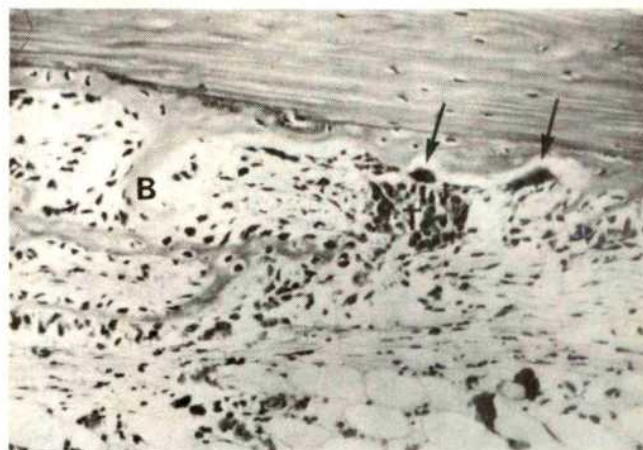
Fig. 2 Section taken 24 h after the injection of VX2 carcinoma. Osteoclasts are prominent (arrowed) near the tumour mass (T), and each is associated with localised bone destruction. Osteoblastic proliferation with new bone formation is also present (B) ($\times 120$).

In the third group (four rabbits) the tumour cell suspension was placed in a diffusion chamber with a 0.45-μm Millipore filter, which was then implanted near the ilium but separated from it by normal muscle. In two of these rabbits a diffusion chamber containing saline was inserted on the contralateral side; the other two had no contralateral operation. The rabbits were killed 8 d later and sections from the left and right pelvis were treated as described previously.

In the fourth group, 4 ml of tumour cell suspension was injected percutaneously into the thigh muscle close to one femur, and the animals were killed 45–55 d later. Sections obtained from the distal and proximal tibiae and the femora were treated as described before. In addition, counts were made of the osteoclasts in each cross section, using a magnification of 200. At least two sections from each site were examined.

Study of the post-mortem sections suggested that there were two distinct mechanisms for the bone destruction. In many instances the bone surface was lined by numerous osteoclasts with local resorption occurring in the lacunae around each osteoclast (Fig. 1). In other sections the bone edge had a trabeculated appearance, suggesting that this was due to osteoclastic lacunae, although no osteoclasts were seen in the lacunae. This occurred with all types of tumour. Osteoclastic proliferation and bone destruction were independent of the associated osteoblastic reaction and new bone production[1]. The osteoclasts were not part of the tumour mass, and usually fibrous reactive stroma filled the space between the edge of the tumour mass and the osteoclasts. Osteoclasts were seen when the associated bone destruction was not gross, but when there were only residual spicules of bone surrounded by tumour cells, osteoclasts were absent.

The first change seen in the rabbits (at 24 h) was osteoclastic proliferation separated from the tumour by a small amount of fibrous stroma (Fig. 2). Surrounding each osteoclast was a small lacuna with bone destruction. Some osteoblastic reaction with new bone formation was also seen at this stage. As the tumour grew, osteoclasts proliferated. As the tumour edge approached the endosteal surface of the cortex, localised destruction occurred in which osteoclasts were prominent. Once the tumour mass actually reached the endosteal surface of the residual cortex, osteoclasts were seen in large lacunae within the cortex. At this stage, there was periosteal new bone formation, associated with the osteoclastic proliferation and cortical bone destruction. When the tumour was injected on to the periosteal surface osteoclastic destruction of the

Fig. 3 Metastasis. Residual trabecula of bone (B) surrounded by tumour (T). Osteoclasts are absent (×96).

cortex, extending from the periosteum to the endosteal surface occurred, the osteoclasts always preceding the tumour.

The osteoclasts disappeared once the tumour had grown sufficiently large to envelop a residual trabecula of bone, but bone destruction continued (Fig. 3). In a few animals osteocytic osteolysis occurred in the tibial cortex, but this accounted for only a minute amount of the bone loss. Occasional infarcts were seen. They usually included bone as well as tumour, and were responsible for only a small amount of the total bone destruction.

Tumour cells implanted in the diffusion chamber remained there as shown by subsequent histological examination, as did the normal muscle between the chamber and the ilium. Osteoclastic proliferation was seen on the periosteal surface and even in the medullary cavity. There was no osteoclastic proliferation on the contralateral side.

Tremendous osteoclastic proliferation occurred in the femur when tumour cells were injected into an adjacent muscle. Their number diminished progressively as the distance from the tumour increased. There was no

Fig. 4 VX2 cell suspension injected into the left thigh, and the number of osteoclasts per cross section of bone counted at various sites. The number progressively diminishes the more distal the section from the tumour. There is no osteoclastic proliferation on the contralateral side.

Number of osteoclasts

osteoclastic proliferation on the contralateral side (Fig. 4).

These experiments suggest that destruction of bone associated with metastatic invasion involves at least two mechanisms, and possibly others of less importance. The initial mechanism, apparently the more important quantitatively, was mediated by osteoclasts. The diffusion chamber experiments suggest that the tumour secreted a humoral factor which stimulated osteoclastic proliferation locally but not on the contralateral side. This was confirmed by the experiments with large amounts of tumour cell suspension implanted in the thigh muscle; the number of osteoclasts progressively decreased distally. Since the contralateral side was unaffected, the results indicate that if this substance enters the circulation, it is rapidly inactivated. Further experiments suggest that prostaglandins are involved[2].

The tumour cells seem to mediate late bone destruction by a mechanism not involving osteoclasts. Osteocytic osteolysis seems of little importance; it was found in the tibial cortex in only a few rabbits and caused only a minute amount of total bone destruction. It was not seen in the human necropsy material. Tumour proliferation may be associated with thrombosis of the vessels and ischaemia, but this was not a feature in the experimental animal or in the necropsy material.

Previous studies have shown that skeletal metastases are associated with both new bone formation and destruction, which occur simultaneously[1]. There seem to be at least two mechanisms for the bone formation. When the tumour provides a suitable matrix for ossification, bone is laid down if osteoprogenator cells are available, for example, metastases from prostatic carcinoma. New bone is formed in all metastases once there is some bone destruction. This is similiar to callus repair after a fracture and may be due to stress on the weakened bone[1]. The present studies suggest that this occurs during osteoclast-mediated bone destruction. Once the destruction was gross and malignant cells surrounded the residual bone, no new bone was formed and no osteoclasts were seen. This may explain the occasional false negatives in skeletal scintigrams when X-rays demonstrate large lytic metastases—at this stage there is no reactive new bone to concentrate the bone-seeking isotope.

C. S. B. GALASKO

Orthopaedic Unit,
Royal Postgraduate Medical School,
London W12 0HS, UK

1 Galasko, C. S. B., *J. Bone Jt Surg.*, 57-B, 353–359 (1975).
2 Galasko, C. S. B., and Bennett, A., *Nature*, 263, 508–510 (1976).

Relationship of bone destruction in skeletal metastases to osteoclast activation and prostaglandins

Two principal mechanisms seem to be responsible for the bone destruction associated with skeletal metastases. The earlier and quantitatively most important mechanism seems to be mediated by osteoclasts stimulated by material produced by the tumour[1]. We now report evidence that this material is, or contains, prostaglandin. There are precedents for such a conclusion. Various tumours contain prostaglandins[2], which resorb bone in organ culture[3]. Prosta-
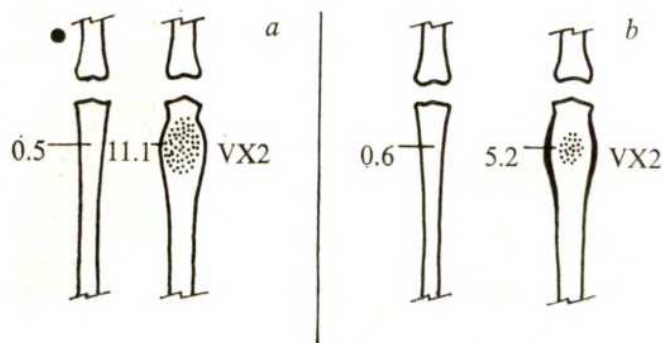
Fig. 1 The effect of indomethacin on osteoclast proliferation near the VX2 carcinoma. There is a marked reduction in the number of osteoclasts (per 100 possible sites) in the tumour-bearing tibiae but no change in the control tibiae. a, Control; b, treated with indomethacin.

glandin formation may explain the growth of dental cysts in the human jaw[4], and prostaglandin synthetase inhibitors reduce tumour-induced osteolysis[5] and hypercalcaemia[6].

The experimental model used is good for investigating skeletal metastases. Approximately 1 g (fresh weight) of VX2 tumour, grown in the thigh muscles of donor New Zealand white rabbits, was finely divided by scissors in 5 ml of 0.15 M sterile NaCl, producing a suspension containing approximately 2.5×10^8 cells per ml. In 16 New Zealand white rabbits 1 ml of cell suspension was injected into the thigh muscle. Eight animals were given indomethacin 7.5 mg per kg body weight per d in their drinking water, starting at intervals varying from 7 d before tumour implantation to immediately after implantation.

The other eight animals, paired with animals receiving indomethacin, served as controls. Each pair received tumour cells from the same donor and was killed simultaneously. Tumours were examined for prostaglandins by homogenising weighed amounts in either Krebs' solution (to indicate synthesising ability) or in 50% ethanol acidified to approximately pH 3 with formic acid (to indicate "basal" levels)[8]. The prostaglandin-like material was extracted and assayed on the rat fundus strip preparation against prostaglandin E_2 (PGE_2)[9]. Part of the extracted Krebs' solution homogenate was characterised by AII solvent system chromatography using paper impregnated with silica gel[10] followed by bioassay.

In 16 other rabbits 1 ml of tumour cell suspension was

Fig. 2 X rays of tibiae removed 31 d after injection of VX2 tumour cell suspension. The tibia on the left is from the control animal, that on the right from the indomethacin-treated animal. There is a marked reduction in the amount of bone destruction in the indomethacin-treated animal.
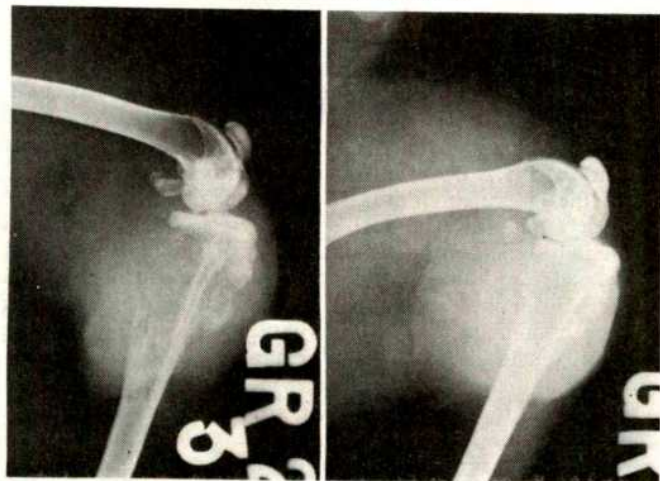
injected into the proximal tibia through a fine drill hole which exactly fitted a bent 18-gauge needle. Half the animals were treated with indomethacin (7.5 mg per kg body weight per d in the drinking water starting 1–7 d before tumour implantation). Indomethacin-treated animals were paired with controls which received tumour from the same donor. Radiographs of the tibiae were taken at intervals and when each pair was killed simultaneously 4–45 d later. Both tibiae were removed, fixed in formalin, decalcified in 5% nitric acid and embedded in wax. At least three histological sections, 1 mm apart, were cut from the proximal part of each tibia, and stained with haematoxylin and eosin. The osteoclasts were counted in each section using a cross-hatched eyepiece graticule and a magnification of 200: an osteoclast was counted if it lay under one of the cross hatches, but not if it lay between two cross hatches. Six hundred randomly selected sites were examined in each section and the number of osteoclasts per 100 sites was calculated.

The animals drank the indomethacin-treated water and the dose of indomethacin required was achieved. In both experimental groups, the lungs at autopsy were filled through the trachea with 15% Indian ink in distilled water containing 1 ml of ammonia per l of solution, removed and fixed in acetic acid–ethanol–formalin–distilled water (2 : 28 : 4 : 66 v/v). The pulmonary metastases, appearing as white areas on a black background, were counted on the external surface and in 1-mm slices through each lobe. The statistical tests included the paired t test.

High concentrations of prostaglandin-like material were extracted from the tumours from animals not receiving indomethacin (Krebs' solution homogenates 4.73 ± 1.06 (s.e.)μg PGE_2 equivalents per g; acid ethanol homogenates 2.4 ± 0.63 μg). The difference between these results is statistically significant ($P < 0.01$). Tumours from indomethacin-treated animals contained significantly less prostaglandin-like material (0.19 ± 0.07 and 0.24 ± 0.15 μg PGE_2 equivalents, Krebs' solution and acid ethanol homogenates respectively, both significantly less than controls, $P < 0.01$). Paper chromatography of extracted Krebs' solution and acid ethanol homogenates (nine and seven tumours, respectively) indicated that all the material was PGE_2. The activity ran fractionally slower than authentic PGE_2, but pure PGE_2 added to the extracts (three experiments) was slowed to the same extent presumably due to extracted impurities.

The number of osteoclasts in the tumour-bearing bone was significantly less in the indomethacin-treated animals (5.2 ± 1.71 osteoclasts per hundred sites) compared with controls (11.1 ± 2.13; $P < 0.05$) (Fig. 1). The numbers were low in the contralateral tumour-free tibiae of indomethacin-treated (0.6 ± 0.09 osteoclasts) and control animals (0.5 ± 0.22 osteoclasts). X rays revealed that in the indomethacin-treated animals bone destruction was markedly diminished but not prevented (Figs 2 and 3).

The number of pulmonary metastases tended to be higher (253 ± 67) in the indomethacin-treated animals than in the controls (189 ± 40), but the difference was not statistically significant ($P > 0.4$).

There is substantial evidence that secretion of prostaglandin by the tumour cells contributes to the activation of osteoclastic bone destruction. The tumour seems to be able to synthesise substantial amounts of PGE_2 as judged by chromatography and bioassay. This supports the conclusion of Voelkel et al. based on radioimmunoassay, the only difference being the amount of PGE_2-like material present. We found 4.73 ± 1.06 μg PGE_2 equivalents per g fresh tissue in our Krebs' solution homogenates. This represents the "basal" amount present plus prostaglandin newly synthesised during homogenisation, less any inactivated. Our "basal" levels are indicated by the acid ethanol homogenates where enzyme activity is inhibited.

Fig. 3 X rays of tibiae removed 45 d after injection of the VX2 cells. The tibia on the left is from the control animal, that on the right from the indomethacin-treated animal. There is a marked reduction in the amount of bone destruction in the latter.

The "basal" values of 2.4 ± 0.63 µg PGE_2 per g are surprisingly high compared with other tissues (for example, human breast[11,12]). Voelkel *et al.* call their level of 0.294 ± 0.051 µg PGE_2 per g fresh weight "content", but since this was obtained after homogenising in a balanced salt solution, it presumably includes newly synthesised material and should be compared with aqueous homogenates. Apart from biological variation, their approximately 16 times lower levels might be due to losses during preparation and extraction (recoveries not stated); differences in time after implantation are unlikely to be important. Our studies confirm the findings of Voelkel *et al.* that feeding indomethacin to the rabbits greatly reduces the amount of extractable tumour prostaglandin. Although they looked at hypercalcaemia whereas we looked at bone destruction, both sets of data strengthen each other.

The reduction of the number of osteoclasts by indomethacin can be explained by the blockade of prostaglandin synthesis which this drug produces[13]. The failure to eliminate all production of osteoclasts and bone destruction might be due either to incomplete inhibition of prostaglandin synthesis, or to the involvement of other osteoclast-activating material (for example, a parathormone-like principle[14]). Furthermore, it is clear that bone destruction can occur without osteoclast involvement in late stages when the tumour cells surround the bone. It therefore seems likely that indomethacin or similar drugs would inhibit both the formation of human bone metastases, and their growth in the early stages. This supports the trial undertaken with benorylate in breast cancer patients with no evidence of bone metastases (personal communication from T. J. Powles), particularly as the synthesis of prostaglandin-like material by human breast tumours may be involved in the development of bone metastases[11,12]. Indomethacin has been used in the treatment of hypercalcaemia due to malignant disease[15] but was not successful in advanced mammary cancer (T. J. Powles, personal communication). These metastases have not been examined histologically, but their radiological findings are consistent with stage II disease[1].

In rats with Walker tumour, aspirin and indomethacin inhibited the hypercalcaemia and osteolysis, demonstrated radiographically, but not examined histologically[5]. There was no significant difference in the soft tissue deposits; the tendency for more pulmonary metastases in our experiments was not statistically significant, and indomethacin treatment did not alter longevity (unpublished data). These results do not support the suggestion[16] that PGE_2 is immunosuppressive and that indomethacin might retard tumour

growth. The formation of soft tissue metastases during treatment with indomethacin may have implications for treatment of bone metastases in human subjects, although it may not be appropriate to compare the effects of a massive injection of tumour cells into rabbits with a slower shedding of cells from human primary tumours. Presumably cells shed slowly from the primary tumour in the presence of indomethacin would not readily develop into skeletal metastases, but it remains to be seen whether they develop elsewhere. Chemotherapy would seem necessary to eradicate the tumour cells, and it might be desirable to treat the patients simultaneously with inhibitors of prostaglandin synthesis.

C. S. B. GALASKO

Orthopaedic Unit,
Royal Postgraduate Medical School,
London W12 0HS, UK

A. BENNETT

Department of Surgery,
King's College Hospital Medical School,
Denmark Hill,
London SE5 8RX, UK

1 Galasko, C. S. B., *Nature*, **263**, 000 (1976).
2 Sykes, J. A. C., and Maddox, I. S., *Nature new Biol.*, **237**, 59–61 (1972).
3 Klein, D. C., and Raisz, L. G., *Endocrinology*, **86**, 1436–1440 (1970).
4 Harris, M., Jenkins, M. V., Bennett, A., and Wills, M. R., *Nature*, **245**, 213–215 (1973).
5 Powles, T. J., Clark, S. A., Easty, D. M., Easty, G. C., and Nevill, A. M., *Br. J. Cancer*, **28**, 316–321 (1973).
6 Voelkel, E. F., Tashjian, A. H., Jr, Franklin, R., Wasserman, E., and Levine, L., *Metabolism*, **24**, 973–986 (1975).
7 Galasko, C. S. B., *J. Bone Jt. Surg.*, **57-B**, 353–359 (1975).
8 Bennett, A., Stamford, I. F., and Unger, W. G., *J. Physiol., Lond.*, **229**, 349 (1973).
9 Unger, W. G., Stamford, I. F., and Bennett, A., *Nature*, **233**, 336–337 (1971).
10 Stamford, I. F., and Unger, W. G., *J. Physiol., Lond.*, **225**, 4P–5P (1972).
11 Bennett, A., McDonald, A. M., Simpson, J. S., and Stamford, I. F., *Lancet*, i, 1218–1220 (1975).
12 Bennett, A., Charlier, E. M., McDonald, A. M., Simpson, J. S., and Stamford, I. F., *Prostaglandins*, **11**, 461–463 (1976).
13 Vane, J. R., *Nature new Biol.*, **231**, 232–235 (1971).
14 Tashjian, A. H., Jr, Voelkel, E. F., Goldhaber, P., and Levine, L., *Fedn Proc.*, **33**, 81–86 (1974).
15 Seyberth, H. W., *et al.*, *New Eng. J. Med.*, **293**, 1278–1283 (1975).
16 Plescia, O. J., Smith, A. H., and Crinwich, K., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 1848–1851 (1975).

Decreased renal prostaglandin catabolism precedes onset of hypertension in the developing spontaneously hypertensive rat

WE have shown in kidneys from normotensive rats, that dramatic age-related changes take place in the catabolism, but not in the biosynthesis, of prostaglandins (PGs)[1]. Although PG catabolic enzymes were active in the late prenatal period, they rose to a peak (60-fold relative to the adult) around 19 d postnatally, dropping to adult levels by day 40 (ref. 1). We have termed this period of peak activity of PG catabolism the 'critical prostaglandin period'[2]. Since the PG catabolising system is principally located in the cortex[3,4], and since cortical maturation is almost entirely a postnatal event (during the first 4–5 weeks) supported by an intrarenal blood flow redistribution[5], the sharp increase in PG catabolism implied an intrinsic requirement by the developing kidney to eliminate locally-formed or plasma PGs. PG catabolism possibly functioned as a protective measure[1] against the potentially harmful effects of the PGs which in this species possess potent vasoconstrictor properties[6]. We theorised that unless inactivated, these compounds could interfere with the redistribution of blood flow to the cortex and consequently with normal corticogenesis. We were therefore interested in investigating the activity of the PG

system in spontaneously hypertensive rats (SHR) at various stages of pre- and early hypertension to compare this activity with age-paired groups of control normotensive rats in an attempt to determine first whether an abnormality in the prostaglandin system exists, second, whether this abnormality appears around the 'critical prostaglandin period' and, third, whether this abnormality develops before or after the onset of hypertension.

Litters (4–7 d) of male SHRs of the Wistar Okamoto–Aoki strain and age-paired normotensive (NR) controls of the Wistar–Kyoto strain were purchased from Laboratory Supplies (Indianapolis) and housed in our animal quarters. Free access to food and water was provided. At various ages one rat from three SHR and three NR litters was killed. The kidneys from each group were combined and homogenised in twenty volumes of buffer (0.05 M KH_2PO_4–$NaOH$, pH 7.4).

PG catabolism was assayed by the radioisotope-dilution method reported previously[1]. Briefly, 0.5 ml homogenate (corresponding to 50 mg wet tissue, approximately 6.5 mg protein) was added to test tubes containing 5,6-$^3H_2$-PGF$_{1\alpha}$ (500,000 d.p.m., New England Nuclear, specific activity 89 Ci mmol^{-1}) containing 0, 1, 5, 10, 30 and 50 μg unlabelled PGF$_{1\alpha}$ and NAD (4 mM final concentration) and tubes were incubated for 10 min at 30 °C. Ethanol (2.5 ml) was added to terminate the incubation and after centrifugation, the supernatant was transferred and evaporated to dryness *in vacuo*. Each sample was assayed qualitatively by radio thin-layer chromatography using chloroform–methanol–acetic acid–water (90:9:1:0.65 v/v) as developing solvent (R_f PGF$_{2\alpha}$ = 0.13; 15K-PGF$_{2\alpha}$ = 0.25, 15K-H$_2$-PGF$_{2\alpha}$ = 0.36) and each product subsequently quantitated by scintillation counting as reported previously[1]. Protein in homogenates was measured by the method of Lowry[9].

PG biosynthetic capacity was measured by mass fragmentography using the stable-isotope dilution technique[7]. An aliquot (10 ml, corresponding to 0.5 g wet tissue, approximately 65 mg protein) of the same homogenate was incubated with excess arachidonic acid (100 μg g^{-1}) 10 min, 37 °C in an oxygen atmosphere. The incubation mixture was acidified to pH 3 with 0.5 N HCl and diethyl ether was added (10 volumes). A mixture of 3,3,4,4-tetradeuterated PGE$_2$, PGF$_{2\alpha}$ and 15-keto-13,14-dihydro-PGF$_{2\alpha}$ (1 μg each) was added followed by a mixture of high specific activity 5,6-tritiated PGE$_1$, PGF$_{1\alpha}$ and 15-keto-13,14-dihydro-PGF$_{1\alpha}$ (500,000 d.p.m. each, approximately 0.7 ng). After extraction, conversion to methyl esters and purification by thin-layer chromatography (R_f PGF$_{2\alpha}$Me = 0.28; PGE$_2$Me = 0.40; 15K-H$_2$-PGF$_{2\alpha}$Me = 0.48), each compound was quantitated as described previously[8].

Table 1 shows the PG biosynthetic and catabolising capacity of

Table 1 Renal biosynthesis and catabolism of PGs by the normotensive (NR) and spontaneously hypertensive (SHR) male rat at various ages

Age		15-PGDH ng per min per mg protein	PGE$_2$ + PGF$_{2\alpha}$ ng per g tissue per 10 min
4 d	NR	76.5	
	SHR	100.1	
17 d	NR	221.6	380
	SHR	124.2	460
49 d	NR	8.9	1,620
	SHR	6.9	2,120
Adult	NR	7.6;6.5	1,520
	SHR	2.9;2.7	2,440

homogenates of kidneys from SHR and NR rats at four different ages. As observed previously[1] 15-PGDH catabolising activity varies immensely with age (about 60-fold around 3 weeks with respect to the adult) whereas the biosynthetic activity does not. Although the combined levels of PGE$_2$ and PGF$_{2\alpha}$ at 17 d are approximately 20–30% of the adult, this difference can be largely explained by the high catabolising activity at 17 d. On the other hand, by comparing the relative capacities of 15-PGDH and the biosynthetic systems between NR and SHR rats it is evident that after 17 d of age the SHR rats possess lower renal catabolic activity. This is also supported by the apparent higher biosynthetic activity of SHR rats (Table 1).

The progressive difference between renal 15-PGDH activity in SHR and NR rats can best be seen in Fig. 1. Renal 15-PGDH activity in male adult SHR is approximately 40% of that found in NR. Before 1 week of age, SHR renal 15-PGDH activity is greater than NR decreasing to adult levels shortly after by 3–6 weeks. On the other hand, although no blood pressure differences are observed between the two types of rats at 3 weeks of age, differences become evident around 7 weeks of age. It is interesting that around this time a spurt in 15-PGDH activity occurs but this quickly returns to adult levels by 10 weeks of age.

The pattern of decreased 15-PGDH activity with age might not be typical of all types of hypertension. Armstrong *et al.*[10] reported similar observations as ours with adult genetic hypertensive rats (GH) of the New Zealand strain but they failed to observe any differences with the Japanese spontaneous hypertensive rat. Unfortunately these authors did not carry out ontogenetic studies with either one of these strains, nor did they provide an indication of the types of PGs formed by these rats.

Our experiments demonstrate that (1) an abnormality in renal PG catabolism (Fig. 1) is evident in SH rats of the Wistar Okamoto–Aoki strain. (2) This abnormality develops in the pre-hypertensive stage reaching the adult level before the appearance of an elevation in blood pressure (3) higher PG biosynthetic capacity is observed in adult SHR than NR, but this reflects the lesser 15-PGDH activity. Thus it is possible that the hypertension that develops around 5–7 weeks of age in SH rats is due to renal 'damage' resulting from the inability to degrade effectively PGs especially during the 'critical prostaglandin period'.

C. R. PACE-ASCIAK

Research Institute,
The Hospital For Sick Children,
Toronto M5G 1X8, Canada

Fig. 1 *a*, Age-dependent difference between the renal PG catabolising activity of spontaneously hypertensive rats (SHR) with respect to normotensive rats (NR); see text for experimental procedure. *b*, Mean arterial blood pressure measurements in SHR and NR rats at three ages. Values represent the mean of two experiments. Rats were anaesthetised with Inactin (100 mg kg^{-1}) and blood pressure was monitored by means of a Statham pressure transducer attached to a cannula in the carotid artery.

[1] Pace-Asciak, C., *J. biol. Chem.*, **250**, 2795–2800 (1975).
[2] Pace-Asciak, C. R., and Rangaraj, G., *J. biol. Chem.*, **251**, 3381–3385 (1976).
[3] Larsson, C. and Änggård, E., *Eur. J. Pharmac.*, **21**, 30–36 (1973).
[4] Pace-Asciak, C., and Domazet, Z., *Biochim. biophys. Acta*, **380**, 338–343 (1975).
[5] Nash, M. A., and Edelman, C. M., Jr, *Nephron*, **11**, 71–90 (1973).
[6] Malik, K. U., and McGiff, J. C., *Circulation Res.*, **36**, 599–609 (1975).
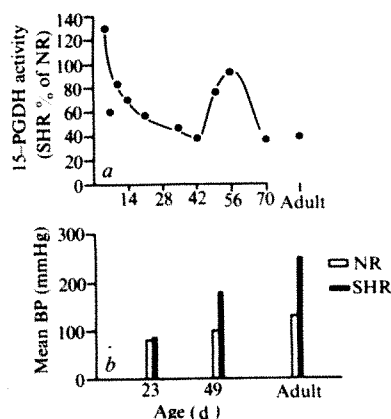
7 Axen, U., Green, K., Horlin, D., and Samuelsson, B., *Biochem. biophys. Res. Commun.*, **45**, 519–525 (1971).
8 Pace-Asciak, C. R., *Advances in Prostaglandin and Thromboxane Research*, 1 (edit. by Samuelsson B., and Paoletti, R.), 35–46 (Raven, New York, 1976).
9 Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J., *J. biol. Chem.*, **193**, 265–275 (1951).
10 Armstrong, J. M., *et al.*, *Nature*, **260**, 582–586 (1976).
11 Pace-Asciak, C. R., *Fedn Proc.*, **35**, 88 (1976).

Transformation of thymocytes by thymus epithelium derived from AKR mice

TYPE C RNA viruses result in the appearance of lymphomas and leukaemias when injected into genetically susceptible strains of mice[1]. The relationship between the murine leukaemia viruses (MuLV) and their target cells in the induction of leukaemia remains, however, incompletely defined. Although MuLV viruses are ubiquitous in the tissues of many mouse strains and are present throughout the life of the animal, they transform specific haemopoietic cell types into leukaemic cells only after a relatively long latent period[2–5]. This suggests that leukaemia is not caused by the virus alone but that a special interaction exists between viruses and target cells in a particular stage of differentiation and maturation.

The target cells of thymotropic MuLV such as radiation leukaemia virus and Gross virus seem to be thymocytes and/or prothymocytes, since the transformed cells have many T-lymphocyte characteristics and they first appear as neoplastic foci in the thymus[6–8]. That differentiation events are important in leukaemogenesis is suggested by changes in the mitogen responsiveness of thymocytes during the latent period before the onset of disease and by change in the relative amounts of cell surface Thy 1 and H–2 antigens in both the spontaneous AKR- and radiation-induced C57BL mouse leukaemia models[9–11]. In both of these systems the incidence of leukaemia is dramatically reduced after thymectomy, which demonstrates the central role played by the thymus[12,13]. The thymic microenvironment may therefore be involved directly or indirectly, through factors released by the epithelium, in the induction of leukaemia by MuLV.

In earlier studies we have used cultures of pure thymus epithelial cells (TE) to simulate the thymic microenvironment *in vitro* to study the induction of differentiation of precursor cells into mature T lymphocytes[14,15]. In the studies reported here, we used TE cultures to investigate the contribution of the thymic microenvironment to leukaemogenesis *in vitro*.

To analyse the role of target cell differentiation as a cofactor in leukaemogenesis we have used lymphoctyes naturally in-

Table 1 Survival of AKR/J mice receiving syngeneic 6-week-old AKR/J thymocytes cultured on thymus epithelium from aged mice

Inducing monolayers	Survival of 6-week old AKR/J recipients†
MEF (AKR/J)*	264 ± 63 d
Kidney epithelium (KE) (4 weeks)	251 ± 69 d
KE (> 24 weeks)	243 ± 74 d
Thymus epithelium (TE)	256 ± 70 d
TE (> 24 weeks)	46.8 ± 6.9 (31–61)

*Mouse embryonic fibroblasts.
†Survivals are reported as the mean ± s.d.
Thymocytes obtained from 6-week-old female AKR/J were cultured together with the inducing cell monolayers in 60 mm Falcon culture dishes (30×10^6 thymocytes per 60-mm dish) containing RPMI 1640 + 15% FCS + 2-mercaptoethanol. After 48 h of culture the monolayers were washed by gentle pipetting with Hanks' balanced salt solution (HBSS). The thymocytes were washed in HBSS and resuspended in phosphate-buffered saline. Thymocytes from each group of inducing monolayers were injected intravenously into 8-week-old syngeneic recipients. 2×10^6 thymocytes were transferred into each animal and each group consisted of 30 mice. The groups were then studied for survival. Only the group receiving cells cultured on TE cells derived from AKR mice over 24 weeks old died after a short latent period with a mean survival of 46.8 ± 6.9 d. The range was between 31 and 61 d. The animals in the control groups did not begin dying of leukaemia until after 150 d.

fected with MuLV as targets for differentiation signals produced by TE. The AKR mouse provides a suitable model since MuLV (Gross type) has been detected in all cell types tested in this strain from before birth[16]. AKR mice develop leukaemia with 90% incidence only after they are 24 weeks old[17]. The AKR model would therefore allow us to study the role of differentiation and maturation of target lymphocytes while avoiding the need to add exogenous MuLV.

The experimental design was to culture TE cells from young non-leukaemic and aged AKR/J mice and to test the ability of these TE cells to transform naturally infected young AKR lymphocytes into neoplastic cells *in vitro*. The technique of TE cell culture has been published in full elsewhere[18]. Briefly, thymuses were removed asceptically, minced, and treated with trypsin–EDTA solution (GIBCO) and collagenase (Worthington Biochem). The cells were grown in monolayer culture in Dulbecco's modified Eagle's medium containing 15% foetal calf serum (FCS). Control monolayers consisted of kidney cells, from young and aged mice, as well as embryonic fibroblasts.

Thymocytes obtained from 6–8-week-old AKR/J mice were cultured together with supporting cell monolayers listed in Table 1 for 48 h. The thymocytes were collected from the culture dishes by gentle pipetting of the monolayer surface with Hanks' balanced salt solution (HBSS). In this way the monolayers are not disrupted and supporting cells are therefore not transferred along with the thymocytes. To ensure this, the culture dishes were each examined microscopically after the thymocytes were removed. The thymocytes were then washed in HBSS and resuspended in phosphate-buffered saline for injection into 8-week-old syngeneic recipients. Thymocytes cultured on TE monolayers from aged (more than 24 weeks old) AKR mice and then transferred into syngeneic recipients killed these animals after a short latent period. The animals died of lymphoma and leukaemia between 31 and 61 d after injection, as evidenced by splenomegaly, lymphadenopathy, and leukaemic nodules in the liver at post-mortem pathology. TE of aged AKR mice were effective in inducing leukaemia whether or not they were derived from the thymuses of overtly leukaemic mice. In contrast, thymocytes cultured on TE obtained from 4-week-old AKR mice did not kill the syngeneic recipients, and these animals died of spontaneous leukaemia only after 24 weeks of age as did uninjected control AKR mice. Also, as shown in Table 1, cultures of embryonic AKR fibroblasts or kidney epithelium from young or old, overtly leukaemic mice were unable to shorten the latent period.

These results indicated that the injection of the donor thymocytes led to the appearance of leukaemia in young syngeneic recipients. It was possible, however, that the leukaemia itself originated in the recipient mice and the leukaemogenic process was only hastened by the injection of the donor thymocytes. It was therefore necessary to ascertain whether true leukaemogenesis of the donor thymocytes had been induced *in vitro*.

We used AKR thymocytes with different alleles of the Thy 1 genetic marker to establish whether the recipient animals died of leukaemia or donor or host cell origin. AKR/J thymocytes, bearing the surface antigen Thy 1.1, that had been cultured on TE cells were injected into 8-week-old AKR/Cu mice whose T lymphocytes bear the Thy 1.2 alloantigen. In every AKR/Cu mouse studied the majority of the leukaemic cells in their spleens were positive for Thy 1.1, showing the disease to be of donor rather than host cell origin (Table 2). When AKR/Cu (Thy 1.2) thymocytes were cultured on TE for 48 h and injected into AKR/J (Thy 1.1) mice, the spleens of the recipients contained Thy 1.2 positive leukaemic cells (Table 2).

The genetic origin of the TE cells whether from AKR/J (Thy 1.1) or AKR/Cu (Thy 1.2) did not determine the *Thy 1* alleles of the leukaemic cells (Table 2). This argues against the possibility that the leukaemia was derived from the old TE cells themselves or from thymocytes that could have originated from the old AKR mice from which the TE cells were obtained. We concluded therefore that the leukaemia cells developing

Table 2 Donor origin of transformed leukaemic cells *in vitro*

| Donor thymocytes | | Inducing TE cells from | | Thy 1 type on leukaemic spleen cells | | |
Strain	Thy 1 Allele	24+-week-old AKR mice Strain	Allele	Recipient	% Thy 1 ±s.e. Thy 1.1	Thy 1.2
AKR/J	Thy 1.1	AKR/J	Thy 1.1	AKR/Cu	83±1.2	8±0.5
AKR/J	Thy 1.1	AKR/Cu	Thy 1.2	AKR/Cu	80±0.9	6±0.5
AKR/J	Thy 1.1	AKR/J	Thy 1.1	AKR/J	71±1.6	5±0.9
AKR/Cu	Thy 1.2	AKR/J	Thy 1.1	AKR/J	9±0.8	79±1.7
AKR/Cu	Thy 1.2	AKR/J	Thy 1.2	AKR/Cu	<3	87±1.4
AKR/Cu	Thy 1.2	AKR/Cu	Thy 1.2	AKR/Cu	<3	78±1.7

Thymocytes obtained from 6-week-old female AKR/J or AKR/Cu mice were cultured together with thymus epithelium from AKR/J or AKR/Cu mice more than 24 weeks old in RPMI 1640+15% foetal calf serum+2-mercaptoethanol. Thymocytes were collected by gently pipetting the monolayers with Hanks' balanced salt solution. The thymocytes were resuspended in phosphate-buffered saline and 5×10⁶ cells were transferred to 8-week-old AKR/J or AKR/Cu mice by intravenous injection. Each group contained 30 animals. Fifteen animals were killed after 30 d and cytotoxicity experiments were performed on individual spleens with anti-Thy 1.1 and anti-Thy 1.2 sera. Single-cell suspensions were prepared and the cell concentration was adjusted to 5×10⁶ viable cells ml⁻¹ with a final concentration of antisera at 1:8. Guinea pig complement was added after 30 min at a concentration of 1:4. Percentage viable cells was determined by Trypan blue exclusion. The results are reported as the means±s.e. of the maximum cytotoxicity of individual spleens.

Also 15 animals from each group were studied for survival. 100% of the mice in each group died after a short latent period. The results show that in each group the leukaemic cells were of donor rather than host origin

in the recipient mice originated from the donor thymocytes which had been incubated with the TE cells. Thus, the TE cells induced leukaemic transformation of young thymocytes naturally infected with MuLV.

Did the cultured thymocytes actually become leukaemic during the *in vitro* culture period? Did the TE cells provide a signal in itself sufficient for leukaemic transformation, or did they merely initiate a process that required further differentiation stimulated *in vivo* by the thymuses of the recipient mice?

To study these questions we performed thymectomy of recipient mice and injected them with donor thymocytes which had been cultured on TE cells. As shown in Table 3, we found that the donor thymocytes could manifest leukaemia even in thymectomised recipients. This finding suggests that the induction of leukaemia was initiated by contact with TE cells cultured *in vitro* and that continued contact with a thymus *in vivo* was unnecessary for leukaemic transformation. Thus 48 h or less of contact with TE cells *in vitro* was sufficient for triggering of leukaemogenesis.

In the present study thymus epithelium from aged AKR/J mice has been shown to initiate directly leukaemic transformation of thymocytes *in vitro*, in the absence of any added MuLV. When TE cells are injected into mice either subcutaneously or under the kidney capsule no tumour is formed. Injection of TE cells does, however, induce leukaemia in the recipient animals. Here the transformed cells are of host rather than donor origin. Mice receiving TE cells from old AKR mice die of lymphoma and leukaemia between 3 and 5 weeks (S. D. Waksal, in preparation).

What is the mechanism of this leukaemogenesis? Other tissues of aged AKR mice such as kidney epithelial cells contain replicating MuLV particles and yet were unable to transform thymocytes *in vitro*. In fact the target lymphocytes themselves already contain similar MuLV. Therefore, leukaemogenesis must involve more than the mere transfer of such viruses.

It has been shown that the thymus has an important role in this disease, since thymectomy up to 5 months of age can markedly inhibit the incidence of the disease in AKR mice[18]. The thymus is responsible for the differentiation and maturation of thymocytes and thymus-derived lymphocytes. The target cell for both normal thymic activity and leukaemogenic transformation seems to be the prothymocyte or thymocyte[8,19,20]. Thus the target cells of the leukaemogenic process might have to be in a critical stage of differentiation to be susceptible to the transforming potential of the MuLV with which they are already infected. The TE cells could activate leukaemia by inducing this necessary stage of differentiation. In support of this concept is our finding that mature lymph node lymphocytes cannot be transformed into leukaemic cells by old TE *in vitro* (S. D. Waksal, unpublished results). Such lymphocytes which have

already passed through the thymic microenvironment are also refractory to further differentiation caused by the TE cells *in vitro* (S. D. Waksal et al., in preparation).

Leukaemic transformation of lymphocytes or their precursors has also been achieved *in vitro* using Abelson virus[21,22]. Here too the state of target cell differentiation seems to be important since the target cells used for transformation are either lipopolysaccharide-stimulated spleen cells or foetal liver cells. The former population contains B lymphocytes non-specifically triggered to differentiate and the latter a population of cells which contain precursors differentiating into B lymphocytes.

It is, however, unlikely that a normal differentiation signal is the sole contribution to leukaemogenesis of the TE cells in our system. TE cells obtained from young AKR mice are capable of inducing the normal differentiation of thymocytes into T lymphocytes but do not provide the leukaemogenic signal (Table 1). Other factors must, therefore, distinguish old from young thymus epithelium. Such factors might include mechanisms which allow the endogenous MuLV to become leukaemogenic within the thymic microenvironment. There may also be a requirement for a second helper virus or viruses which replicate in the TE cells to produce leukaemogenesis in combination with the Gross-type MuLV. TE cells from aged AKR mice have significantly higher levels of replicating MuLV than TE cells from young mice. This has been observed using

Table 3 Effect of thymectomy on survival of AKR mice receiving syngeneic 6-week-old AKR/J thymocytes transformed by TE cells *in vitro**

Inducing monolayer (AKR/J)	Survival ATx recipient†
Kidney epithelium (4 weeks)	> 285 d
Kidney epithelium (>24 weeks)	> 330 d
TE (4 weeks)	> 330 d
TE (> 24 weeks)	44±6.9 d (32–53)

*Thymocytes from 6-week-old female AKR/J mice were cultured together with the inducing monolayers for 48 h. The thymocytes were collected by gentle pipetting of the monolayer surfaces and 5×10⁶ cells were transferred into groups containing 25 syngeneic 8-week-old recipients. Each recipient had been thymectomised at 4 weeks of age. Only 1 out of 25 mice which had been adult thymectomised but received no transferred cells developed lymphoma with latent period of 187 days. AKR/J mice which had been thymectomised and received thymocytes cultured together with TE monolayers from AKR/J mice over 24 weeks old died of lymphoma after a short latent period (44±6.9 d) with splenomegaly and lymphadenopathy. Only 2 animals in the control adult thymectomised groups died of leukaemia. The other control mice were still alive beyond days listed in Table 3.

†Survival of the last group is reported as the mean survival±s d. with a range of 32–53 d.

fluorescent anti-MuLV. These higher levels of virus may be critical in the development of leukaemia. These viruses might include the AKR xenotropic virus[23], which has been shown to appear only in TE cells derived from AKR mice older than 24 weeks.

The TE would then serve a dual role in the aged AKR; to trigger the cell into a susceptible state of differentiation and to supply the additional factors needed to initiate leukaemogenesis. Leukaemogenesis therefore seems to involve two or three distinct phases: (1) the infection with MuLV; (2) the necessary state of target-cell differentiation; and/or (3) additional unknown factors present in TE cells of aged AKR mice necessary to trigger leukaemic transformation. At present it is not possible to distinguish between the second and third possibilities. The long latent period needed for leukaemogenesis in the intact mouse might be required to generate such additional factors as the AKR xenotropic virus in the thymic microenvironment.

The *in vitro* system presented here makes feasible an analysis of the interactions of these factors in leukaemogenesis.

SAMUEL D. WAKSAL

Immunology Branch,
National Cancer Institute,
National Institutes of Health,
Bethesda, Maryland 25014

S. SMOLINSKY
I. R. COHEN
M. FELDMAN

Weizmann Institute of Science,
Rehovot, Israel

1 Rich, M. A., and Siegler, R., *A. Rev. Microbiol.*, 21, 529–572 (1967).
2 Todaro, G. J., and Heubner, R. J., *Proc. natn. Acad. Sci. U.S.A.*, 69, 1009–1015 (1972).
3 Gross, L., *Proc. Soc. exp. Biol. Med.*, 76, 27–32 (1951).
4 Rauscher, F. J., *J. natn. Cancer Inst.*, 29, 515–543 (1962).
5 Kaplan, H. S., *Cancer Res.*, 27, 1325–1340 (1967).
6 Kaplan, H. S., *Ser. Haematol.*, 7, 94–163 (1974).
7 Gross, L., *Oncogenic Viruses* (Pergamon, Oxford, 1970).
8 DeCleve, A., Trevis, Weismann, I. L., Lieberman, M., and Kaplan, H. S., *Cancer Res.*, 35, 3585–3595 (1975).
9 Waksal, S. D., Smolinsky, S., Cohen, I. R., St Pierre, R. L., and Feldman, M., *Adv. exp. Biol. Med.*, 66, 141–146 (1976).
10 Zatz, M. M., *J. Immun.*, 115, 1168–1170 (1975).
11 Haran-Ghara, N., *Adv. exp. Biol. Med.*, 66, 135–170 (1976).
12 McEndy, D. P., Boone, M. D., and Furth, J., *Cancer Res.*, 4, 377–383 (1944).
13 Kaplan, H. S., *J. natn. Cancer Inst.*, 11, 83–90 (1950).
14 Waksal, S. D., *et al.*, *Ann. N.Y. Acad Sci.*, 249, 492–498 (1975).
15 Waksal, S. D., Cohen, I. R., Waksal, H. W., and St Pierre, R. L., *Fedn Proc.*, 33, 736 (1974).
16 Rowe, W. P., and Pincus, T., *J. exp. Med.*, 135, 429–436 (1972).
17 Furth, J., Seibold, A. R., and Rathbone, R. R., *Am. J. Cancer*, 9, 521–532 (1933).
18 Waksal, S. D., Latham-Griffin, E. A., and St Pierre, R. L., *Immun. Commun.* (in the press).
19 Miller, J. F. A. P., *Lancet*, ii, 748–751 (1961).
20 Jankovic, B. D., Waksman, B. H., and Arnason, B. G., *J. exp. Med.*, 116, 159–176 (1962).
21 Raschke, W. C., Ralph, P., Watson, J., Sklar, M., and Coon, H., *J. natn. Cancer Inst.*, 54, 1249–1253 (1975).
22 Rosenberg, N., Baltimore, D., and Sher, C. D., *Proc. natn. Acad. Sci. U.S.A.*, 72, 1932–1936 (1975).
23 Lieber, M. M., Sherr, C. J., and Todaro, G. J., *Int. J. Cancer*, 13, 5–12 (1974).

Hormone-dependent haematopoiesis in fibroblast-transformation ossicles

SUBCUTANEOUS transplantation of coarse powders of acid-insoluble matrix of the shafts of long bones of rat to allogeneic recipients results consistently in the formation of new bone through endochondral ossification[1]. Haematopoietic bone marrow develops in the newly formed ossicle in adolescent rats[2,3] and this experimental model is amenable to an investigation of hormonal control of initiation of haematopoiesis. Erythropoiesis is dependent on pituitary hormones[4,5] since hypophysectomy resulted in anaemia in all animal species which have been studied. Hypophysec-
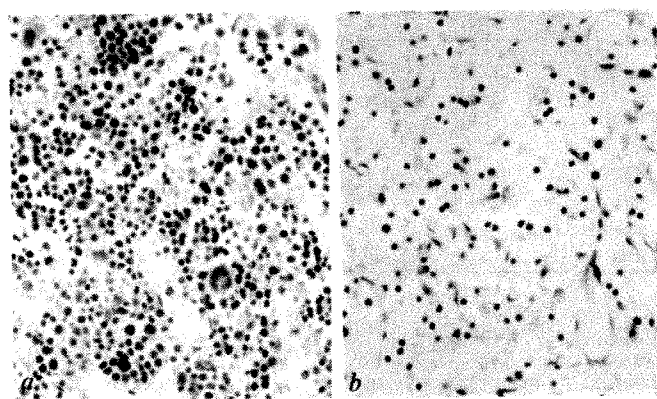
Fig. 1 Photomicrographs of histological sections of transformation ossicles from control (*a*) and hypophysectomised (*b*) rats stained with haematoxylin and eosin. *a*, Haematopoietic cells on day 26 in a control rat. *b*, Scattered basophilic cells in a hypophysectomised rat. Hypophysectomy on day 0 (×250).

tomy caused a prompt regression of many stem-cell erythro-blastic leukaemias in rat[6]. We now report the profound influence of hypophysectomy on *de novo* erythropoiesis in matrix-induced transformation ossicles.

Dehydrated diaphyseal shafts of rat femur and tibia were comminuted and sieved. Powders with particle sizes of 74–420 μm were demineralised with 0.5 M HCl and extracted with water, ethanol and ether as described in ref. 1. This preparation is designated as 'transformant of fibroblasts (TF). Male rats, age 35–37 d, of Long–Evans strain were anaesthetised with ether. In sterile conditions TF (25–30 mg) was transplanted as a compact deposit into a subcutaneous pocket prepared by blunt dissection. The incision was closed with a metallic clip. As previous studies[2] have established the superiority of the thoracic site over that of the abdominal site for erythropoiesis, transplants were confined exclusively to the former anatomical location. In each rat there were two bilaterally symmetrical sites in the upper thoracic region. The day of operation was denoted day 0. Hypophysectomy was performed by D. Otway of the Hormone Assay Laboratory, Chicago, on day 0 unless otherwise indicated in the tables. At the time of collection of the transformation ossicles, the rats were anaesthetised and exsanguinated by cardiac puncture with a 19-gauge hypodermic needle followed by decapitation. The ossicles were dissected out and weighed. The weights of testis and spleen were recorded to confirm the completeness of hypophysectomy. The body weights of all the animals were recorded three times a week. The hypophysectomised animals lost weight whereas the control rats gained 3–4 g per day.

The histological sequences in response to subcutaneous implantations have been described in detail[2,3] and consist of the following stages in the control rats. On day 1 a planoconvex, button-like plaque was formed; it is a conglomerate of TF, fibrin network and polymorphonuclear leukocytes. On day 3 the plaque consisted of TF and invading fibroblasts. By day 4 there was close-range matrix–membrane interactions[3] and chondroblasts emerged on day 5 for the first time. Chondrocytes were abundant on days 7 and 8. With the incursion of capillaries on day 9 there began chondrolysis and osteogenesis. By day 12 an osseous conglomerate with cavernous spaces had formed. On day 14 extravascular islands of dark-staining haemocytoblasts were evident. With progressive differentiation the entire ossicle was filled with haematopoietic bone marrow replete with cells in the erythrocytic and granulocytic series and with megakaryocytes by days 23–25 (Fig.

Table 1 Influence of hypophysectomy on alkaline phosphatase and ^{59}Fe incorporation in transformation ossicles, femoral bone marrow and spleen

| | Alkaline phosphatase U g^{-1} Transformation ossicles | | Transformation ossicles | | ^{59}Fe incorporation (c.p.m. per mg tissue) Bone marrow | | Spleen | |
Day	Control	Hypox	Control	Hypox	Control	Hypox	Control	Hypox
10	60.0±5	38.7±3*	—	—	1,671±198	442±120*	259±47	28±9*
21	23.6±2	60.3±4*	77±35	<1	1,756±147	971±113*	486±43	20±3*
25	29.8±3	45.6±2*	334±49	1±0.3	2,541±131	553±95*	1,030±138	30±9*
28	29.8±2	41.7±3*	579±80	9±0.4				

Alkaline phosphatase and ^{59}Fe incorporation into protein-bound haem was determined as described earlier[1]. All numbers represent mean±s.e. of 8 observations.
*Significant difference from control.

1a). Collection was routinely carried out on days 27–31 in our experiments, as haematopoiesis became fully functional at that time.

Bone formation was assayed by measuring the activity in the transplantation ossicle of alkaline phosphatase, an enzyme known to be intimately associated with ossification[7]. Hypophysectomy (hypox) delayed but did not prevent the formation of the osseous conglomerate. The maximal concentration of alkaline phosphatase, ∼ 60 U g^{-1}, was similar (Table 1) in the ossicles of intact and hypox hosts. But the peak of enzyme activity in intact rats was attained on day 10 whereas in hypox rats, the concentration of alkaline phosphatase in the ossicles was a maximum (Table 1) on day 21.

Hypophysectomy profoundly inhibited the development of haematopoiesis. On day 27, the transplantation ossicles in normal hosts were dark red with an abundance of maturing blood cells (Fig. 1a). In contrast, the ossicles in hypox rats were pale pink in the gross on day 27 and the marrow consisted exclusively of scattered individual basophilic cells (Fig. 1b), presumably haemocytoblasts; maturation of blood cells was absent. The incorporation of ^{59}Fe into protein-bound haem, an indicator of erythropoiesis, was barely detected in transformation ossicles of hypox rats and was significantly lower than the controls in the femoral bone marrow (Table 1). The spleen of hypox rats was tiny and pale pink in the gross; the spleen of control rats was larger and had a dusky red colour (Table 2). The ^{59}Fe incorporation in the spleen was significantly depressed following hypophysectomy (Table 1).

In another series of experiments the day of hypophysectomy was varied to determine if a delay in removal of the pituitary would improve the ^{59}Fe incorporation in transformation ossicle, bone marrow and spleen It was found that hypophysectomy 9 to 26 d before collection resulted in a profound decrease in erythropoiesis as revealed by ^{59}Fe incorporation (Table 2).

The results demonstrate that the onset of haematopoiesis

Table 2 Decline of ^{59}Fe incorporation in transformation ossicles, femoral bone marrow and spleen following hypophysectomy

| Hypophysectomy day | ^{59}Fe incorporation (c.p.m. mg^1) | | |
	Ossicle	Bone marrow	Spleen
	Hypophysectomised rats		
0	2±0.4*	179±50*	10±1*
8	1±0.4*	165±45*	10±1*
12	4±2*	96±39*	8±1*
22	23±6*	297±6*	14±2*
	Control rats; no operation		
—	408±111	3247±169	484±95

Control and hypophysectomised rats were littermates. Autopsy was performed 26–31 d after transplantation of bone matrix. Hypophysectomy was performed on the days indicated. ^{59}Fe incorporation into protein-bound haem was determined as described[1]. All values represent mean±s.e. of 8 observations.
*Significant difference from control.

is critically dependent on pituitary hormones. The availability of a hormone-dependent differentiating haematopoietic system in post-foetal life makes it possible to integrate the factors regulating hormone-dependent and hormone-independent haematopoiesis.

This work was supported by grants from the American Cancer Society, the Jane Coffin Childs Memorial Fund for Medical Research, the National Foundation March of Dimes and the NIH.

A. H. REDDI
CHARLES B. HUGGINS

*The Ben May Laboratory for Cancer Research,
University of Chicago,
Chicago, Illinois 60637*

1 Reddi, A. H., and Huggins, C. B , *Proc. natn. Acad. Sci U S A.*, **69**, 1601–1605 (1972).
2 Reddi, A. H., and Huggins, C. B., *Proc. natn. Acad. Sci U.S A.*, **72**, 2212–2216 (1975).
3 Reddi, A. H., and Anderson, W. A., *J. Cell Biol.*, **69**, 557–572 (1976)
4 Crafts, R. C., and Meineke, H. A , *Ann. N. Y. Acad. Sci.*, **77**, 501–517 (1959).
5 Krantz, S B., and Jacobson, L. O , *Erythropoietin and the Regulation of Erythropoiesis*, 158–164 (University of Chicago, Chicago, 1970).
6 Huggins, C B , and Oka, H., *Cancer Res.*, **32**, 239–242 (1972).
7 Robison, R , *Biochem. J.*, **17**, 286–293 (1923)

Depletion of L-cell sterol depresses endocytosis

THE engulfment of droplets of fluid, from the surrounding medium (endocytosis or pinocytosis), represents a phenomenon of the plasma membrane of many mammalian cells[1-3]. The physiological significance of the process is not well understood although circumstantial evidence seems to suggest that it is necessary for normal functions of certain cell types. Steinman *et al.*[4] have established the qualitative and quantitative aspects of endocytosis and have demonstrated that a variety of cells directly interiorise soluble proteins (for example, horseradish peroxidase) by invagination of the plasma membrane without previous binding of the proteins to cell-surface receptors. Cholesterol is a prominent constituent of the plasma membrane of mammalian cells, and its role in many complex functions of the cell membrane is just beginning to be understood. It buffers the 'fluidity' of the plasma membrane, increasing the viscosity at high temperatures and impeding the transition into the gel state at reduced temperatures[5-7]. Through its action on membrane fluidity, or possibly more directly, it affects specific functions of the membrane such as ion transport and enzyme activities[8]. With these manifold effects of cholesterol on membrane structure and function in mind, we investigated the effect of altered sterol concentration on the ability of cultured cells to carry out endocytosis, to clarify further the role of cholesterol in membrane-dependent cytological functions.

The procedures we used to investigate the role of membrane-associated sterol in endocytosis are based on our finding that certain oxygenated derivatives of cholesterol (for example, 25-hydroxycholesterol and 7-ketocholesterol) are potent inhibi-

Fig. 1 Rate of endocytosis in L cells after treatment with various inhibitors of sterol synthesis. The culture medium contained 25-hydroxycholesterol 1 μg ml^{-1} (△), 20α-hydroxycholesterol, 1 μg ml^{-1} (▲), 7-ketocholesterol 1 μg ml^{-1} (■), 6-ketocholestanol 1 μg ml^{-1} (×) or cholesterol 100 μg ml^{-1} (●). Control cultures contained 50 μl 5% bovine serum albumin (BSA) solution used to suspend the sterols (○). The cultures were incubated for 16 h at 37 °C. Each point represents the mean ± s.e of 3–5 cultures.

tors of sterol synthesis in cultured cells[9]. The inhibitory sterols specifically depress the activity of 3-hydroxy-3-methylglutaryl-(HMG)-CoA reductase, the regulatory enzyme in the sterol synthesis pathway. Treatment of L cell (mouse fibroblast) cultures with these inhibitors for periods longer than 1 d reduces the sterol (desmosterol) content of the cells and their plasma membranes to about one-half of the original level (A.A.K. and H.W.C., unpublished and ref. 10). Cells which are depleted of sterol do not grow or divide unless an exogenous, non-inhibitory sterol (cholesterol or desmosterol) or a sterol precursor (mevalonate) is supplied[10].

L cells were grown in Waymouth's MB 752/1 medium supplemented as described previously[9]. The cells were seeded in 35-mm plastic dishes of 2-ml portions containing a total of 0.7 × 10⁶ to 2 × 10⁶ cells depending on the type of experiment. Different inhibitors of sterol synthesis (6-ketocholestanol, 7-ketocholesterol, 20α-hydroxycholesterol, or 25-hydroxycholesterol) were added to the medium (2 ml) in 50 μl of a 5% solution of crystallised bovine serum albumin (BSA) containing 10% ethanol[9] so that the final concentration of sterol was 1 μg ml^{-1}. Control cultures received 50 μl of the albumin–ethanol solution.

Cellular endocytosis was assayed as described by Steinman *et al.*[4]. Briefly, L cells were incubated for various periods of time with sterols and/or their precursors and were then reincubated for 1–3 h with horseradish peroxidase (Sigma, type II) dissolved in fresh culture medium (1 mg ml^{-1}). The peroxidase-containing medium was removed by suction, the cultures were gently washed five times with phosphate-buffered saline, and then reincubated for 50 min at 37 °C in Waymouth MB 752/1 medium to remove all non-ingested peroxidase. The cells were then lysed by incubation with an aqueous solution of Triton X-100 (0.05% v/v) for 2 h at room temperature. The amount of peroxidase in the lysate was determined as described by Steinman *et al.*[4]. These authors have shown that: there are no cell-surface receptors for the probe (peroxidase); the rate of endocytosis increases linearly with temperature; endocytosis is dependent on ATP

generated by glycolysis and the respiratory chain; and the process is independent of the cell cycle.

Figure 1 shows the effects of incubating various inhibitors of sterol synthesis with L cells for 16 h on the ability of the cells to carry out endocytosis. 20α-hydroxycholesterol and 25-hydroxycholesterol depressed the rate of endocytosis by 42 and 54% respectively, whereas 6-ketocholestanol and 7-ketocholesterol were less active. The relative potencies of these sterols as suppressors of endocytosis were similar to their reported potencies as inhibitors of sterol synthesis[9]. Purified cholesterol at a high dose (100 μg ml^{-1}) had no effect on endocytosis. The rate of uptake of peroxidase in the control cultures averaged 167 ng per mg cellular protein per h. This corresponds approximately to 0.0049% of the administered peroxidase per 10⁶ cells h^{-1}, a value which is slightly higher than the corresponding range of 0.0032–0.0035% reported by Steinman *et al.*[4] in another subline of L cells.

The effect of 25-hydroxycholesterol on endocytosis as a function of time during which the inhibitor was incubated with the cultures is shown in Fig. 2. Endocytosis was quantitatively affected as early as 6 h after addition of 25-hydroxycholesterol. After 12 h the level of endocytosis had reached a minimum level which remained unchanged for the following 36 h. The slope of the curve in Fig. 2 appears to be somewhat steeper than the decline in the concentration of membrane sterol in other similar experiments and was preceded by the disappearance of HMG-CoA reductase activity and the loss of sterol synthesis (ref. 9 and A.A.K. and H.W.C., unpublished). The effect of incubating L cells for 16 h with various concentrations of 25-hydroxycholesterol is shown in Fig. 3. Endocytosis was affected by 25-hydroxycholesterol at a concentration as low as 0.05 μg ml^{-1} (10^{-7} M) and maximal depression of endocytosis was reached at concentrations of 0.5 μg ml^{-1}.

L cells, in which HMG-CoA reductase activity is depressed by 25-hydroxycholesterol, are able to synthesise sterols from exogenous mevalonate, which is the product of the reaction catalysed by the reductase. In similar conditions they are not able to synthesise sterols from acetate or HMG[9,10]. Table 1 shows that the addition of exogenous mevalonate or cholesterol resulted in the maintainance of a normal level of endocytosis in the presence of 25-hydroxycholesterol. In other experiments (not shown) acetate or HMG did not relieve the inhibitory effect of 25-hydroxycholesterol. The recovery of endocytosis with meva-

Fig. 2 Time course of inhibition of endocytosis by 25-hydroxycholesterol. The culture medium contained 5% foetal calf serum (●), 25-hydroxycholesterol 1 μg ml^{-1} in 50 μl 5% BSA solution (▲) or 50 μl BSA solution (○). After the incubation periods indicated, rates of endocytosis over a 2-h period were determined. Values shown represent the mean ± s.e. for 3 cultures.
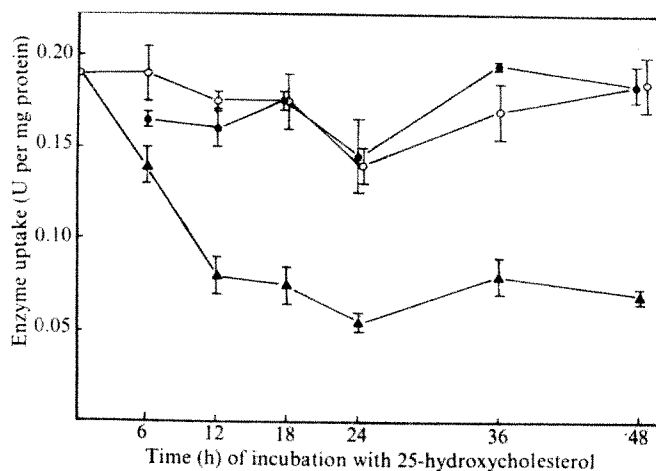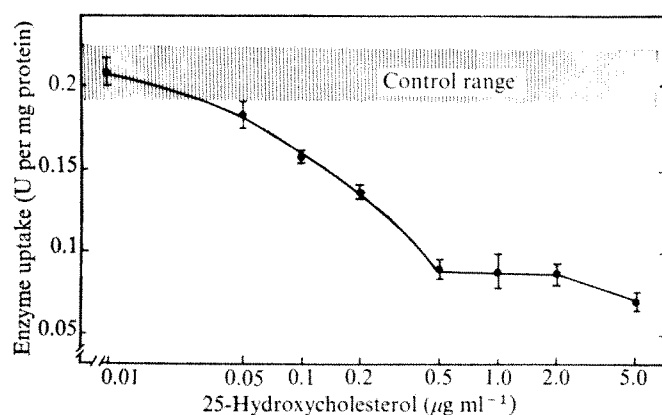
Fig. 3 Effect of incubating L cells with varying concentrations of 25-hydroxycholesterol on the rate of endocytosis. 25-hydroxycholesterol was added to the culture medium in 50 μl 5% BSA solution. After incubation for 16 h with the inhibitor, rates of endocytosis were determined over a 2-h period. Values shown represent the mean ±s.e. for 3 cultures.

Ionate supports our previous findings (refs 9 and 10 and A.A.K. and H.W.C., unpublished) that the inhibitory sterols do not have effects other than those resulting from the suppression of HMG-CoA reductase and sterol synthesis.

The possibility that incubation of L cells with the inhibitory sterols may have depressed the level of intracellular ATP was examined by comparing the concentration of ATP in cultures incubated with or without 25-hydroxycholesterol. Using the luciferin–luciferase assay[11,12], we did not detect any significant differences in the concentration of ATP between control cultures and cultures treated with 25-hydroxycholesterol for as long as 40 h. As the cultures approached confluency the level of ATP dropped in both groups. A diminished ATP level in confluent cell cultures has been reported by others[4].

These results, indicating that cells which are deficient in sterol are unable to carry out endocytosis at normal rates, emphasise the important functional role of cholesterol in the plasma membrane. Dianzani *et al.*[13] have demonstrated that increased amounts of cholesterol in the plasma membrane of macrophages depresses phagocytosis—that is, endocytosis of particulate material. If phagocytosis in macrophages and endocytosis in cultured L cells are accomplished by the same basic mechanism, then our results, together with their observations, indicate that the concentration of cholesterol in mammalian cell membranes must be maintained within a fairly narrow range to ensure their

proper biological functioning. Our results may also be related to certain observations of altered functioning of fibroblasts from patients with familial hypercholesterolaemia. Since fibroblasts from these patients cannot take up low density lipoprotein (LDL) at a normal rate[14] it would be of interest to determine whether or not there is any abnormal concentration of cholesterol in the plasma membranes of these cells. Our observations also suggest an alternative interpretation of data published by Brown and Goldstein[15] which showed a diminished uptake of LDL in normal human fibroblasts that had been preincubated with 25-hydroxycholesterol or LDL. Since both surface-bound and internalised LDL were included in the measurement of LDL 'binding' it could be argued that the effect of 25-hydroxycholesterol was to diminish endocytosis rather than to reduce the number of LDL receptors as originally suggested by Brown and Goldstein.

It is established that cholesterol has an important role in the regulation of the activity of membrane-localised Na^+/K^+ ATPase in mycoplasma[9] and in mammalian cells[8,16,17]. Thus, changes in endocytosis and in ion transport both may contribute to the inhibition of DNA synthesis that we have observed in L cells and in lectin-stimulated lymphocytes treated with inhibitors of sterol synthesis (ref. 18 and A.A.K. and H.W.C., unpublished). Whether the continuous uptake of solutes and membrane material is necessary for the maintenance of normal cell functions needs further clarification. The present study suggests, however, that cholesterol is a critical functional as well as structural unit in the plasma membrane and that its depletion results in the modification of specific membrane-dependent cellular functions.

This study was supported by the NCI. We thank Mrs Jan Marshall for assistance.

HANS-JORG HEINIGER
ANDREW A. KANDUTSCH
HARRY W. CHEN

*The Jackson Laboratory
Bar Harbor, Maine 04609*

1 Lewis, W. H., *Bull. Johns Hopkins Hosp.*, **49**, 17–36 (1931).
2 Cohn, Z. A., *J. exp. Med.*, **134**, 557–571 (1966).
3 Holter, H., *Int. Rev. Cytol.*, **8**, 481–504 (1951).
4 Steinman, R. M., Silver, J. M., and Cohn, Z. A., *J. Cell Biol.*, **63**, 949–969 (1974).
5 Kimelberg, H. K., and Paphadjopoulos, D., *J. biol. Chem.*, **249**, 1071–1080 (1974).
6 DeGeir, J., Mandersloot, J. G., and VanDeenen, L. L. M., *Biochim. biophys. Acta*, **173**, 143–145 (1959).
7 Bieri, V. G., and Wallach, D. F. H., *Biochim. biophys. Acta*, **389**, 413–427 (1975).
8 Rottem, S., Yashouv, J., Ne'eman, Z., and Razin, S., *Biochim. biophys. Acta*, **173**, 143–145 (1959).
9 Kandutsch, A. A., and Chen, H. W., *J. biol. Chem.*, **248**, 8408–8417 (1973); **249**, 6057–6061 (1973); *J. cell. Physiol.*, **85**, 415–424 (1975).
10 Chen, H. W., Kandutsch, A. A., and Waymouth, C., *Nature*, **251**, 419–421 (1974).
11 Stanley, P. E., and Williams, S. G., *Analyt. Biochem.*, **29** 381–392 (1969).
12 Hammerstedt, R. H., *Analyt. Biochem.*, **52**, 449–455 (1973).
13 Dianzani, M. U., Torrielli, M. V., Canuto, R. A., Garcea, R., and Feo, F., *J. Path.*, **118**, 193–199 (1976).
14 Stein, O. W., Weinstein, D. B., Stein, Y., and Steinberg, D., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 14–18 (1976).
15 Brown, M. S., and Goldstein, J. L., *Cell*, **6**, 307–316 (1975).
16 Papahadjopoulos, D., *J. theor. Biol.*, **43**, 329–337 (1974).
17 Kimelberg, H. K., *Biochim. biophys. Acta*, **413**, 143–156 (1975).
18 Chen, H. W., Heiniger, H. J., and Kandutsch, A. A., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 1950–1954 (1975).

Table 1 Rate of endocytosis in L cells

Additions*	Peroxidase uptake† (ng per mg cellular protein per h)	% Control	*t* test
Control	242.1±29.5 (7)	100	
25-Hydroxycholesterol (1 μg ml⁻¹)	109.0±26.9 (7)	45	P<0.01
25-Hydroxycholesterol (1 μg ml⁻¹)+ mevalonic acid (0.5 mg ml⁻¹)	229.2±20.8 (2)	95	NS
25-Hydroxycholesterol (1 μg ml⁻¹)+ cholesterol (100 μg ml⁻¹)	250.0±27.8 (2)	103	NS

*Cells were cultured for 36 h in the presence of the listed compounds. †The rate values (mean±s.e. for 3–4 cultures) were based on measurements obtained between the first and second hour of incubation with horseradish peroxidase. Numbers in parentheses indicate the number of experiments. The *t* test results indicate the probability of a significant difference between control and experimental values.

Duplication of biochemical changes of Huntington's chorea by intrastriatal injections of glutamic and kainic acids

SYSTEMIC administration of large doses of L-glutamate to immature animals causes degeneration of neurones in the retina and the arcuate nucleus. Olney *et al.*[1] reported that intracerebral injections of the more potent excitatory analogues of glutamate, in particular kainic acid, caused rapid degeneration of certain dendritic and somal structures in the injected area. Presumably the degeneration is related to the ability to bind the "excitotoxic" agent, with necrosis resulting from a sustained increase in membrane permeability.[3] ³H-kainic acid binds more strongly to striatal tissue than to any other area of rat

brain tested[2]. This lends some support to Spencer's theory[3]—based on antagonism of cortical excitation of striatal neurones by diethyl glutamate—that the massive corticostriatal pathway is glutaminergic. If Huntington's chorea were an excitotoxic phenomenon resulting from chronic overstimulation of glutamate receptors, striatal neurones in particular would be affected because of heavy glutaminergic input. This hypothesis is not readily tested by simple measurement of glutamate in choreic tissue because glutamate is a precursor of γ-aminobutyric acid (GABA) and is also involved in other metabolic activities. As one potential test of this hypothesis, we have injected 1 μl of either kainic acid or L-glutamic acid dissolved in isotonic saline directly into the extrapyramidal nuclei of rats. Injection into the striatum produced local enzymatic changes duplicating those reported in Huntington's chorea.

The sites of injection and the coordinates by the Konig and Klippel atlas[4] were: caudate, AP 8.38, L 2.8, DV +0.2; ventrolateral thalamus, AP 4.75, L 2.8, DV +1.4; and substantia nigra, AP 2.18, L 2.1, DV −2.2. Injections were into 300-g male Wistar rats under sodium pentobarbital anaesthesia.

The extrapyramidal system has neurones containing dopamine, acetylcholine and GABA. The state of these neurones in extrapyramidal disorders such as Huntington's chorea and Parkinson's disease is reflected by the level of the enzymes for transmitter synthesis[5]. Therefore we measured the levels of tyrosine hydroxylase (TH), choline acetyltransferase (CAT) and glutamic acid decarboxylase (GAD) in the rat brain extrapyramidal nuclei after the injections.

After 2 or 3 d, the rats were killed by cervical fracture, the brains were removed rapidly, and the desired nuclei were dissected and homogenised in 0.25 M sucrose. Aliquots of these sucrose homogenates were used for enzyme determinations by the radioactive methods used in the study of enzyme activities in human brain tissue[5]. Protein was determined by the Folin–Ciocalteau method.

The biochemical results are shown in Tables 1 and 2. After injections of 5 nmol of kainic acid (Table 1), the largest changes were observed with striatal injections. There were large losses of GAD and CAT (to 34% and 36%, respectively) in the striatum itself, and there was a smaller decrease of GAD in the substantia nigra (to 66%). There was also a significant increase (to 128%) in striatal TH. With comparable thalamic injections, there were no significant enzyme changes in the substantia nigra and caudate, and only a mild drop in thalamic GAD. Nigral injections also produced no significant striatal changes, although there was a slight drop in nigral TH.

When the amount of kainic acid injected into the caudate was reduced to 2.5 nmol, the effect was reduced, and 1 nmol had no effect. A dose of 50 nmol of L-glutamic acid produced the equivalent effect of 2.5 nmol of kainic acid (Table 2).

Table 1 Some enzyme activities on the side injected with 5 nmol of kainic acid as a percentage of those on the control side

Area injected	Area analysed		
Enzyme	Striatum	Thalamus	Substantia nigra
Caudate (6)			
GAD	34 ± 3 %*	107 ± 16 %	66 ± 14 %†
CAT	36 ± 8 %*	113 ± 16 %	119 ± 21 %
TH	128 ± 9 %†	91 ± 12 %	100 ± 14 %
Thalamus (3)			
GAD	90 ± 15 %	72 ± 11 %†	96 ± 6 %
CAT	109 ± 12 %	114 ± 15 %	—
TH	94 ± 11 %	92 %14 %	—
Substantia nigra (4)			
GAD	107 ± 9 %	106 ± 9 %	125 ± 27 %
CAT	90 ± 10 %	71 ± 8 %†	107 ± 19 %
TH	87 ± 14 %	70 ± 20 %	73 ± 9 %†

Control values were the same as found in uninjected animals and averaged for the striatum (in μm per h per 100 mg protein): 15.5 for GAD; 31.3 for CAT and 0.175 for TH. Protein content was not significantly different in the injected as compared to the control samples. Data given as mean ± s.e., number of animals in parentheses.
*$P < 0.001$.
†$P < 0.05$.

Table 2 Enzyme levels in striatum on injected side as percentage of those on contralateral side

Material injected	CAT	GAD	TH
Kainic acid, 2.5 nmol	81 ± 4 %*	84 ± 6 %*	111 ± 8 %
Kainic acid, 1 nmol	100 ± 8 %	91 ± 7 %	113 ± 12 %
L-glutamic, 50 nmol	82 ± 5 %*	83 ± 4 %*	114 ± 6 %

Figures are mean ± s.e. Four rats were used in each case.
*$P < 0.05$.

The reduction in striatal CAT, and striatal and nigral GAD, with sparing or enhancement of TH, parallels the biochemical findings in Huntington's chorea[5–7]. Thus, these excitotoxic injections produced a biochemical model of the disease. They also produced marked physical manifestations.

Rats injected with 5 nmol of kainic acid into the left caudate almost immediately developed diarrhoea and stopped grooming. By the third day, they were bleeding from their nose and urinary tract. When injections were made in the thalamus there was similar, though lesser bleeding, and loss of grooming. Injections into the substantia nigra did not cause bleeding but did cause the mouth and eyes to water. Rats given the injections in either the left caudate or left substantia nigra held their heads tilted markedly to the left with their bodies hunched. Rotation was not prominent, however, except when injections were made into the thalamus: then each rat, on awaking from the anaesthesia 30–60 min after injection, rolled over and over slowly to the left while the tail was held stiffly erect and rotated clockwise like a propeller and the head was tilted to the left. After 2 or 3 h rolling ceased and the animal alternately stood erect and pirouetted, also clockwise. Subsequently, the rat repeatedly circled the plastic cage in a clockwise direction, rubbing against the wall and with the head still tilted to the left. These movements had stopped by the next day.

Rats injected with 1 or 2.5 nmol of kainic acid or with 50 nmol of L-glutamic acid into the caudate showed no dramatic effects.

With the 5-nmol injections of kainic acid, two of eight caudate-injected rats died between the second and third day, as did three of six thalamus-injected rats. There were no deaths observed in the substantia nigra-injected rats. All three rats that received this dose of kainic acid in the globus pallidus died within 24 h.

The injections were always unilateral and no behavioural tests were carried out. To establish whether the excitotoxic model of Huntington's chorea truly parallels the disease, further exploration of the motor and behavioural effects is needed.

The excitotoxic hypothesis of Huntington's chorea might also help to explain the phenomenon of tardive dyskinesia. This motor disorder is characterised by choreiform movements which are hard to distinguish from Huntington's chorea and which appear after long term therapy with dopaminergic blockers. The dopaminergic nigro-striatal tract is believed to be inhibitory towards striatal interneurones, and administration of neuroleptics apparently releases cholinergic interneurones[8,9]. This removal of inhibition could result in normal excitatory input from the cortex, possibly of a glutaminergic nature, becoming excitotoxic and gradually destroying striatal interneurones.

E. G. McGeer
P. L. McGeer

Kinsmen Laboratory of Neurological Research,
Department of Psychiatry,
University of British Columbia,
Vancouver, British Columbia V6T 1W5,
Canada

1 Olney, J. W., Sharpe, L. G., and de Gubareff, T., *Neurosci. Abst.*, 5, 371 (1975).
2 Simon, J. R., Contrera, J. F., and Kuhar, M. J., *J. Neurochem.*, 26, 141–147 (1976).
3 Spencer, H. J., *Brain Res.*, 102, 91–101 (1976).
4 Konig, J. F. R., and Klippel, R. A., *The Rat Brain* (Williams and Wilkins, Baltimore, 1963).
5 McGeer, P. L., and McGeer, E. G., *J. Neurochem.*, 26, 65–76 (1976).
6 Bird, E. D., and Iversen, L. L., *Brain*, 97, 457–472 (1974).
7 Stahl, W. L., and Swanson, P. D., *Neurology*, 24, 813–819 (1974).
8 McGeer, P. L., Grewaal, D. S., and McGeer, E. G., *Brain Res.*, 80, 211–217 (1974).
9 Stadler, H., Lloyd, K. G., and Bartholini, G., *Naunyn Schmeidebergs Arch. Pharmac.*, 283, 129–134 (1974).

Stereoselective effects of the potentially neuroleptic rigid spiro amines

WE HAVE reported the synthesis and pharmacological effects of a novel series of tetracyclic spiro compounds structurally related to the tricyclic antidepressants[1]. In contrast to the common tricyclic antidepressants, the terminal amino group of the spiro compounds has restricted mobility. As expected, some of these compounds have antidepressant properties in model systems, for example they are potent blockers of the neuronal uptake of noradrenaline in the central nervous system. But the compounds containing a chlorine atom in the tricyclic moiety have a neuroleptic profile in animal tests. 3-Chloro-10,11-dihydro-N,N-dimethylspiro-[5H-dibenzo[a,d]cycloheptene-5,1'-cyclohex-2'-ene]-4'-amine, structure (I), is particularly interesting in this respect. The cyclohexene ring in this structure is located perpendicularly to the tricyclic ring system and as a consequence the amino group can occupy a position either *cis* or *trans* to the chlorine-containing part of the tricyclic ring system. Furthermore, the *cis* and *trans* isomers can be resolved into enantiomers as the molecule contains two chiral centres (marked with asterisks in structure (I)). Several studies have yielded re-

(I)
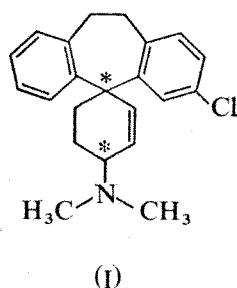
sults to favour the view that neuroleptic drugs act as dopamine (DA) antagonists in the brain[2,3] and that the DA receptor blockade is rather well correlated with their antipsychotic efficiency[4-6]. The neuroleptics seem to block chiefly postsynaptic DA receptors, particularly in striatal[7] and mesolimbic[8] DA areas of the brain. The rigid nature of the isomeric tetracyclic spiro amines of structure (I), and the possibility of stereospecific pharmacological effects make these compounds interesting tools for an investigation of the structural requirements for DA receptor blockade and for studies on DA receptor topology.

The *cis* and *trans* isomers have now been separated in the following way. A substance, melting point 245–246 °C (hydrochloride), prepared essentially as described previously[1], was converted to the corresponding oxalate. Fractional crystallisation of this oxalate gave partial separation. Conversion to the hydrochlorides and renewed crystallisation resulted in the pure isomers. These are referred to as the α isomer (melting point 257–259 °C (hydrochloride)), and the β isomer (melting point 226–228 °C (hydrochloride)).

The separation was monitored by nuclear magnetic resonance spectroscopy. Comparison of the spectra of the pure isomers with that of the substance used in our previous

Table 1 Effects on apomorphine stereotypies in the rat

Compounds	Reg. no*	ED_{50} (μmol kg^{-1})
α + β	A 02056	40
α	A 23622	> 54
β	A 23623	10
(+)-α	A 23885	> 54
(−)-α	A 23886	> 54
(−)-β	A 23887	4
(+)-β	A 23888	> 44
Chlorpromazine		14

*Astra's internal code number.
Apomorphine (2 mg kg^{-1}) was administered intraperitoneally to male Sprague–Dawley rats (150–200 g). The test compounds were injected intraperitoneally 30 min before apomorphine in five doses (six animals per dose). The stereotypies were scored according to Costall and Naylor[11]. ED_{50} refers to the dose of the test compounds which reduces the stereotypies by 50% 20 min after apomorphine administration.

work[1] showed that the latter consisted of a mixture of about equal parts of α and β isomers.

The α and β isomers have been resolved using the optically active mandelic acids and di-p-toluoyltartaric acids. The antipodes of the two isomers have specific rotations of 148° (hydrochloride in ethanol) and 204° (maleate in chloroform), respectively.

The isomers were evaluated for their ability to inhibit apomorphine-induced stereotypies in the rat; this is a selective and sensitive test for neuroleptic activity[9]. Apomorphine stereotypies seem to be a result of DA receptor stimulation in the brain, which makes this test an *in vivo* technique for measuring DA receptor blockade[10].

The results of the apomorphine test show differences in activity between the isomers (Table 1). The *cis–trans* mixture of the spiro amine (I) is a potent blocker in this test. The pharmacological activity resides chiefly in the β isomer which is about as potent as chlorpromazine. Studies on the enantiomers revealed a further difference in activity. Only one of the enantiomers, (−)-β, was found to be active in this test, being three or four times as potent as chlorpromazine. Further evidence for the specificity of this enantiomer as a blocker of the DA receptor was provided by DA turnover studies in the rat. It is well known that neuroleptic drugs accelerate DA turnover, probably due to an increase in DA neuronal activity as a result of DA receptor blockade[2-6]. Of the two pairs of enantiomers, only the (−)-β isomer increased DA turnover in the brain.

In studies of thioxanthenes, the *cis* isomers, were found to be considerably more potent than the corresponding *trans* isomers in various tests for neuroleptic activity[12,13]. It may therefore be postulated that the active tetracyclic spiro amine (−)-β has the *cis* configuration. The stereochemistry of the isomers is being investigated by X-ray diffraction methods.

We conclude that there are strict stereochemical requirements for a potent and selective blocker of the DA receptor. The compounds discussed here apparently fulfil these requirements.

BERNT CARNMALM
LARS JOHANSSON
STEN RÄMSBY
NILS E. STJERNSTRÖM

Department of Organic Chemistry,

SVANTE B. ROSS
SVEN-OVE OGREN

Department of Pharmacology,
Research and Development Laboratories,
Astra Läkemedel AB,
S-151 85 Södertälje, Sweden

1 Carnmalm, B., *et al.*, *J. med. Chem.*, 17, 65–72 (1974).

2 Carlsson, A , and Lindqvist, M , *Acta pharmac. Tox.*, 20, 140–144 (1963)
3 Iversen, L. L., *Science*, 188, 1084–1089 (1975).
4 Nyback, H., Borzecki, Z., and Sedvall, G., *Eur J. Pharmac.*, 4, 395–403 (1968)
5 Carlsson, A., Persson, T., Roos, B. E., and Walinder, J , *J. Neur. Transmiss.*, 33, 83–90 (1972)
6 Snyder, S. H , Banerjee, S. P., Yamamura, H. I , and Greenberg, D., *Science*, 184, 1243–1253 (1974).
7 Kebabian, J. W., Petzold, G. L., and Greengard, P., *Proc natn. Acad. Sci. U.S.A.*, 69, 2145–2149 (1972).
8 Horn, A. S., Cuello, A. C., and Miller, R. J., *J. Neurochem.*, 22, 265–270 (1974).
9 Janssen, P. A. J., Niemegeers, C. J. E , Schellekens, K H. L., and Lenaerts, F. M., *Arzneimittel. Forsch*, 17, 841–854 (1967).
10 Ernst, A. M., *Psychopharmacologia*, 10, 316–323 (1967).
11 Costall, B., and Naylor, R. J., *Eur. J. Pharmac.*, 24, 8–24 (1973).
12 Møller-Nielsen, I., *et al.*, *Acta pharmac. Tox.*, 33, 353–362 (1973).
13 Miller, R. J., Horn, A. S., and Iversen, L. L., *Molec. Pharmac.*, 10, 759–766 (1974)

Ca²⁺-dependent arrest of cilia without uncoupling epithelial cells

IN CERTAIN ciliated epithelia, for example, mussel gill lateral cell epithelium, metachronally coordinated ciliary beat can be arrested systematically, either after nervous[1] or local stimulation, for example by laser microinjury[2,3]. In the latter case, the velocity of spread of arrest, its extent and decremental character suggest that the response depends on electrotonic coupling of the gill cells[3]. The lateral cells are coupled by extensive septate junctions and small gap junctions[4], more basally located, which may act in a manner homologous to electrical synapses of nerve and muscle cells[4-7]. Although intracellular microelectrode recordings of lateral cell depolarisation accompanying arrest after branchial nerve stimulation have been made in *Mytilus*[1], there are no reports of microelectrode studies of cell coupling in gill cells. We have used experimentally-induced spreading arrest to monitor the state of coupling in the lateral cell epithelium of a freshwater mussel (for example, *Elliptio complanatus*).

Lateral cell cilia are arrested by an increase in cytoplasmic calcium ion concentration. Arrest does not spread past a focus of local damage in *Mytilus* gill in the absence of Ca²⁺ in the bathing medium[3]. Furthermore, when gill filaments from freshwater mussels are placed in 12.5 mM CaCl₂, the cilia beat metachronally; when a divalent cation ionophore A23187 is added, arrest occurs; arrest does not occur by ionophore in the presence of Na, K⁺ or Mg²⁺, without Ca²⁺ (ref. 8).

An increase in local Ca²⁺ concentration causes coupled epithelial cells of *Chironomus* salivary gland to uncouple[8], but Rose and Loewenstein[14] have shown that Ca²⁺ influx into a cell is often quenched by various cell components, so that Ca²⁺ concentration in the cytoplasm may change only minimally, if at all, in parts of the cell distant from the site of influx. It may therefore be that even when Ca²⁺ is entering a cell at the apical surface, the effective concentration of Ca²⁺ at the cell junctions remains low and the cells remain coupled. This seems to be a necessary conclusion in many circumstances, for example, during cardiac
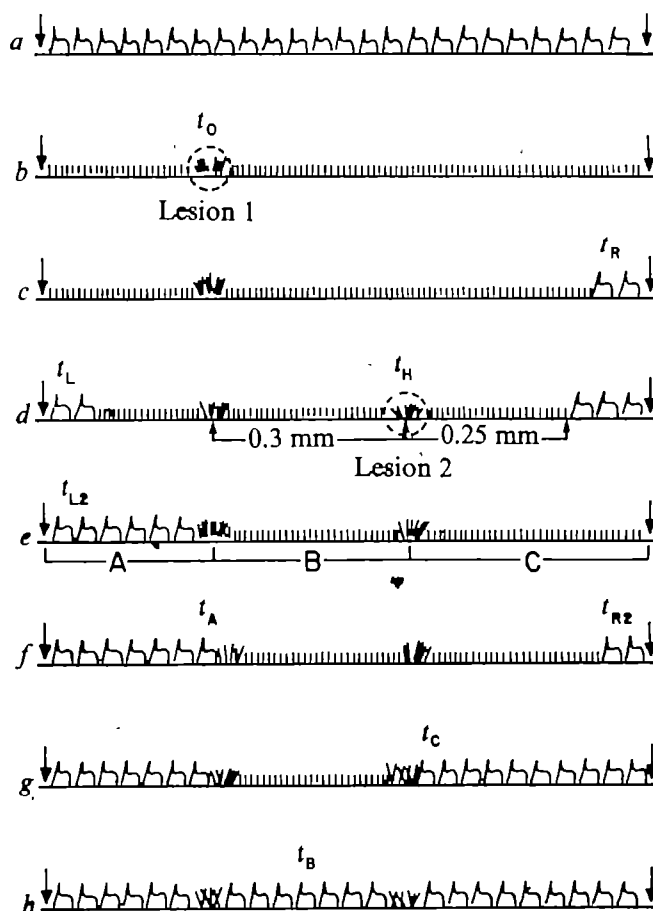
Fig. 1 Design of summation experiment. A gill filament beating with good metachronal waves (η-like structures) selected (a). At t = 0, a laser hit (lesion no. 1) made; within a few seconds along a defined field (length about 1 mm, arrows) metachronism ceased (b). The initial recovery to either side of the field was measured (t_R and t_L) (c). The gill was repositioned so that a second hit (lesion no. 2) could be made about 0.25 mm away from a returning wavefront (d). At t_H, this hit divided the gill filament into regions A, B and C (e). The effect on recovery was measured by the reappearance of metachronism to either side of the field (t_{L2} and t_{R2}). The geometry varied slightly in individual experiments. Usually region A recovered completely (t_A) (f); then region C (t_C) (g) and finally region B recovered (t_B) (h) with a time course given by equation (3) (see text). Cells damaged by the hits never recovered.

muscle contraction, but as far as we know it has not been demonstrated precisely with any system. We report here that mussel gill cells remain coupled during ciliary arrest.

Gills of freshwater mussels (Carolina Biological Supply Co., Gladstone, Oregon) were excised and placed in 12.5 mM CaCl₂; they usually remained fully active for many

Table 1 Typical arrest and recovery parameters after laser microinjury

Experiment no.	R	L	H	L₂	L₂-H	R₂	R₂-H	A	C	0.63A+(C-H)	B
1	23	32	40	40	0	80	40	93	125	144	140
2	20	15	26	26	0	52	26	58	75	86	87
3	8	7	10	10	0	20	10	12	115	113	115
4	13	5	14	14	0	25	11	79	52	88	97
5	10	8	11	11	0	33	22	20	60	62	60
6	55	50	56	56	0	90	34	117	105	123	155
7	61	60	64	110	46	99	35	150	105	135	129
8	25	70	40	—	—	58	18	105	115	141	160
9	44	75	50	—	—	95	45	90	115	122	115
10	7	17	9	—	—	15	6	28	24	33	31
Average*	24.7	34.4	27.8	—	—	54.3	26.5	71.1	78.5	95.5	103.3
s.d.	22.9	24.6	23.8	—	—	36.0	18.5	39.4	42.5	36.1	41.3

*31 experiments, t test: A compared with B, P < 0.001; C compared with B, P < 0.01; (0.63A+C-H) compared with B, P > 0.3.

hours with almost 100% metachronism. Small pieces of active gill were placed on a depression slide containing three or four drops of 1% blue dye FDC No. 1 in $CaCl_2$, for laser microdisruption as described before[2,3,11]. A pulsed ruby laser microbeam was focused near the base of the beating cilia at a selected point to produce a focal lesion of about $30\,\mu m$ diameter (Fig. 1a and b). The timing of all subsequent events (Table 1) was measured from the moment of microbeam irradiation (t_0).

The lesion induced arrest of ciliary beat which spread, as expected, along the gill filament up to about 1 mm on either side of the hit (Fig. 1b)[2,3]. Characteristically, metachronism was first reactivated in the cells most distant from the hit. The time that metachronism reappeared in a defined field to either side of the lesion (t_R and t_L) was recorded (Fig. 1c and d). Between the original lesion and the returning metachronal wave cilia were still arrested. The gill was then moved so that the laser microbeam was focused near the centre of the arrested field to one side of the lesion, approximately 0.25 mm away (10 cell lengths) from the returning metachronal wavefront, so as to produce a second lesion at this point (Fig. 1d). The gill filament was thus divided into three regions: A, B, and C, where B lies between the two lesions, and C is beyond the second lesion (Fig. 1e).

In about 90% of experiments, the second lesion reintroduced arrest in region C (Fig. 1e). Experimental failure probably reflects our inability to focus the laser accurately and the consequent variability of microbeam strength at lateral cell sites. This also accounts for much of the variability in the data, shorter times probably reflecting weaker hits. We measured the time after the second hit when metachronism first reappeared in region C (t_{R2-H}) (Fig. 1f). Typical data are given in Table 1; recovery was delayed on the average by 26.5 s (31 experiments).

To produce such delay, the stimulus produced by the second lesion must apparently have spread through an average of 10 lateral cells whose cilia are arrested, before reaching the reactivating cells where we observed secondary arrest. This supports the conclusion that the cells with arrested cilia were still coupled, but does not rule out an alternative possibility that cells (for example, the intermediate cells) adjacent to the column of arrested lateral cells provided the coupling pathways, being coupled to one another and to the reactivating cells at the time of the second hit.

In contrast, the second lesion reintroduced arrest in region A, beyond the first lesion, in only 20% of experiments; that is, recovery (t_{L2-H}; Table 1) is not delayed (Fig. 1e). We interpret this to mean that lesion No. 1 usually destroyed the physical integrity of the gill filament, so that the column of cells beyond the lesion to the left was uncoupled from the column of cells to the right. Therefore, the stimulus produced by the second lesion should have affected the cilia of regions B and C but not those of region A. If this is correct, the proability is greatly reduced that coupling that affects the lateral cell metachronism can occur through an alternate pathway adjacent to the lateral cells.

We have demonstrated conclusively that cells with arrested cilia are still coupled, at least as defined by the ability to respond to a second distant stimulus, by measuring the times for complete recovery in regions, A, B and C (t_A, t_B and t_C, respectively; Fig. 1f, g and h). If arrested cells are uncoupled, we would expect recovery in region B to have been unaffected by the second lesion so that t_A and t_B would have been approximately equal. This was not the case (Table 1). If arrested cells are coupled, then region B would have been affected by the second hit, so that $t_B > t_A$ because of the induced recovery delay. Two simple explanations consistent with coupling were that (1) t_B and t_C are equal or (2) that region B would have recovered as if the hits were additive.

The first laser hit presumably produces a stimulus that enables sufficient Ca^{2+} from the bathing fluid to enter the apical surfaces of cells near the lesion to induce ciliary arrest in sequential fashion. Our results demonstrate that the pathway of coupling remains functional after ciliary arrest, although it is readily interrupted by a break in epithelial integrity. Our best explanation of the summation effect in the region between the two hits is that (1) the second hit produces further Ca^{2+} influx into the coupled chain of cells with arrested cilia and that (2) the recovery times measured depend on Ca^{2+} being pumped out of the cells at a constant rate such that when free Ca^{2+} falls below a critical concentration, metachronism reappears. If regions B and C are of approximately equal lengths and the stimulus is exponentially decremental with a decay constant of about 0.3mm, then we would expect that

$$t_B = t_A + \text{(effect of second lesion)} \qquad (1)$$

and that the effect of the second lesion is related to t_C in that

$$t_C - t_H = \text{(effect of second lesion)} + 0.37 t_A \qquad (2)$$

so that (3)

$$t_B = 0.63 t_A + t_C - t_H \qquad (3)$$

Both in individual experiments and on average ($P > 0.3$) equation (3) seems empirically valid (Table 1). This explanation supports and extends findings of Rose and Loewenstein[10] in that it requires the actions of Ca^{2+} in these conditions to be restricted largely to those in the immediate locale of the point of influx of the ion into the cytoplasmic matrix of the cell, that is to the cilia and not to the junctions.

PETER SATIR
WILLIAM REED
DAVID I. WOLF

Department of Physiology-Anatomy,
University of California,
Berkeley, California 94720

1 Murakami, A., and Takahashi, K., *Nature*, 257, 48–49 (1975).
2 Satir, P., Fong, I., and Goldstein, S. F., *Acta Protozool.*, 11, 287–290 (1972).
3 Motokawa, T., and Satir, P., *J. Cell Biol.*, 66, 377–391 (1975).
4 Gilula, N. B., and Satir, P., *J. Cell Biol.*, 51, 869–872 (1971).
5 Furshpan, E. J., and Potter, D. D., *J. Physiol., Lond.*, 145, 289–324 (1959).
6 Loewenstein, W. R., *Ann. N.Y. Acad. Sci.*, 137, 441–472 (1966).
7 Satir, P., and Gilula, N. B., *A. Rev. Ent.*, 18, 143–166 (1973).
8 Satir, P., *Science*, 190, 586–588 (1975).
9 Rose, B., and Loewenstein, W. R., *Nature*, 254, 250–252 (1975).
10 Rose, B., and Loewenstein, W. R., *Science*, 190, 1204 (1975).
11 Goldstein, S. F., *J. exp. Biol.*, 51, 431–441 (1971).

Nervous inhibition of corpora allata by photoperiod in *Pyrrhocoris apterus*

PHOTOPERIOD governs reproduction in many insect species with adult diapause. The nervous and neuroendocrine systems mediate the impact of the environment on ovarian functions. It has been suggested that the daylengths promoting diapause prevent secretion of the activation hormone (most probably a complex of neurohormones) from the pars intercerebralis (PI) of the brain. Without this hormone the corpora allata (CA) do not secrete the gonadotropic hormone and reproduction is arrested[1-3]. It has been shown, however, that other physiological conditions such as pregnancy[4], fasting[5], or lack of mating[6] produce an inhibitory effect of the brain on the CA through the nervi allati. The inhibition is probably mediated by a nervous centre situated close to the neurosecretory cells of the PI. We show here that the photoperiod that promotes diapause arrests reproduction by a double mechanism, the
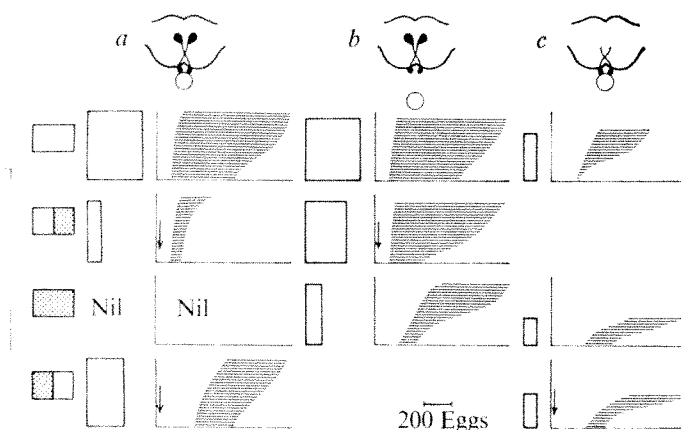
Fig. 1 Schematic representation of the effect of denervation of corpora allata and extirpation of pars intercerebralis on reproductive activity. The first column of oblongs indicates photoperiods. Open, long day (LD 18 : 6); shaded, short day (LD 12 : 12). The three series of oblongs and graphs represent: *a*, sham operations (*n* = 20, 13, 35, 30); *b*, transections of nervi allati (*n* = 28, 17, 18); *c*, extirpations of pars intercerebralis (*n* = 10, 25, 44). Oblongs represent fecundity; width, eggs per female; height, % ovipositing females. Shaded area denotes length of oviposition period in days; horizontal axis, % ovipositing females (length of axis = 100%), vertical axis, time (length of axis = 50 d). Arrows, change of photoperiod.

suppression of humoral stimulation and the inhibition of the CA through the nerves. The effect of photoperiod is mediated by the pars intercerebralis.

We have studied the role of the nervi allati and centres in the PI in the regulation of the secretory activity of the CA in females of *Pyrrhocoris apterus*, under long- and short-day regimes. This species exhibits a facultative diapause controlled very regularly by the photoperiod; long daylengths have a stimulatory effect on the CA and reproduction, whereas short daylengths are inhibitory. This enables us to study both the stimulatory and inhibitory mechanisms regulating the activity of the CA. Experimental females were reared individually in Petri dishes on linden-seed at 26±2 °C and in long-day (18:6 h of light : darkness) conditions or short-day (12:12) conditions respectively. Transection of the nervi allati and extirpation of the PI were performed through an incision of the neck membrane[7]. Since size is not a reliable sign of CA activity in *P. apterus*[8], the activity of the CA was evaluated by the rate of oviposition.

In females kept permanently under long-day conditions the denervation of the CA did not affect reproduction and therefore the activity of the CA. The same operation disturbed the inhibitory effect of short daylengths however. The control (sham-operated) females stopped oviposition soon after the transfer from long- to short-day conditions, or did not begin to oviposit when reared from the eggs at short daylengths. In the first case, transection of the nervi allati resulted in a considerable prolongation of the oviposition period and increased fecundity. In the second case, the females with denervated CA had at the least low reproductive activity (Fig. 1*a*, *b*). This demonstrates that the inhibitory signal of short daylengths arrives at the CA through the nervi allati. It is interesting that transection of the nervi allati in diapausing *Leptinotarsa decemlineata*[1] or *Anacridium aegyptium*[3] had no effect. Transection of the nervi allato-suboesophagealis slightly activated the CA in *L. decemlineata*, but this activation was not sufficient to induce oviposition[1].

Extirpation of the PI in *P. apterus* resulted in low reproductive activity regardless of the photoperiod used (Fig. 1*a*, *c*). Since sham-operated females were very fecund under long-day conditions and they did reproduce under short-day conditions, the PI seems to mediate the optimal stimulation of reproduction

by long daylengths and also its complete inhibition by short daylengths.

In long-day conditions the PI stimulated reproduction not only through the CA but also by other pathways; the treatment of females without PI with a juvenile hormone analogue accelerated the maturation of oocytes in the first reproductive cycle, but it did not increase the overall fecundity[9]. As denervation of the CA did not affect reproduction in long-day conditions we assume that the PI stimulates the CA by way of the haemolymph. In short-day conditions, however, the PI seems to inhibit the CA through the nerves; denervation of the CA stimulated reproduction in a similar way to the extirpation of the PI. Similar rates of reproduction in both these cases also indicate that at short daylengths the PI does not exert the stimulatory effect as observed at long daylengths. The reason why denervated CA are more active in long-day females than in short-day females[10] can probably be explained by the lack of humoral stimulation from the PI in short-day conditions.

The findings described enable us to propose a hypothetical model of photoperiodic regulation of reproduction in *P. apterus* (Fig. 2). According to this model the PI contains both stimulatory and inhibitory centres. The photoperiod regulates the relative activities of these two centres. In long-day conditions activation of the stimulatory centre predominates which results in a highly active CA and reproduction. Conversely, in short-day conditions the activity of the inhibitory centre prevails, the CA are inactivated and females do not oviposit at all (Fig. 2*a*). The denervation of the CA does not affect the reproductive activity of long-day females for the stimulation of the CA is humoral. The same operation, however, induces oviposition in short-day females, because the inhibitory signal does not arrive at the CA through the sectioned nervi allati. Reproductive activity is low as the action of the stimulatory centre is suppressed by short daylengths (Fig. 2*b*). Both the stimulatory and inhibitory centres are removed by the extirpation of the PI. This results in a low baseline reproductive activity for both photoperiods (Fig. 2*c*).

The common belief that the inactivity of the CA and the arrest of reproduction are caused by the diapause-promoting photoperiod only through an absence of the activation hormone

Fig. 2 Hypothetical model of photoperiodic regulation of reproduction in *P. apterus*. *a*, Intact females; *b*, females with nervi allati transected; *c*, females with pars intercerebralis extirpated. Large upper circle, brain; smaller circle below, corpora allata. Small circles inside the brain, regulatory centres; open, active stimulatory centre; solid, active inhibitory centre; shaded, inactive stimulatory or inhibitory centre. Open arrows, stimulatory action; solid arrow, inhibitory action. Effect on ovaries: +, high reproductive activity; (+), low reproductive activity; −, no reproductive activity.
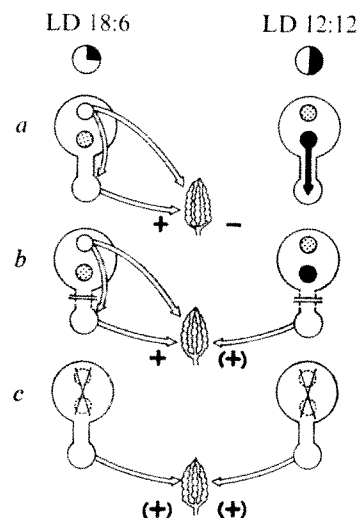
cannot thus apply to all insects. In some species at least, reproductive inactivity is assured by a double mechanism—not only by the absence of humoral stimulation but also by the inhibition of the CA through the nerves. Such evidence has been obtained in *P. apterus*[9] and in *Locusta migratoria*[11]. A more detailed description of the present results will be given elsewhere.

M. HODKOVÁ

Institute of Entomology,
Czechoslovak Academy of Sciences,
Na Folimance 5, 120 00 Praha 2,
Czechoslovakia

1 De Wilde, J., and de Boer, S. A., *J. Insect Physiol.*, 15, 661–675 (1969).
2 Girardie, A., and Granier, S., *J. Insect Physiol.*, 19, 2341–2358 (1973).
3 Girardie, A., Moulins, M., and Girardie, J., *J. Insect Physiol.*, 20, 2261–2275 (1974).
4 Engelmann, F., and Lüscher, M., *Verh. Deut. Zool. Ges. Hamburg, 1956*, 215–220 (1957).
5 Johansson, A. S., *Nytt. Mag. Zool.*, 7, 3–132 (1958).
6 Friedel, T., *J. Insect Physiol.*, 20, 717–733 (1974).
7 Sláma, K., *J. Insect Physiol.*, 10, 283–303 (1964).
8 Hodková, M., *Vest. Cs. spol. zool.* (in the press).
9 Hodková, M., thesis, Czechoslovak Academy of Sciences, Praha (1975).
10 Hodková, M., *J. Insect Physiol.* (in the press).
11 Darjo, A., *J. Insect Physiol.*, 22, 347–355 (1976).

TRH potentiates excitatory actions of acetylcholine on cerebral cortical neurones

ANTAGONISM of barbiturate-induced sleeping time in experimental animals[1] is one of the interesting pharmacological properties of thyrotropin-releasing hormone (TRH; pyroglutamyl-histidyl-proline amide) which are independent of its effects on endocrine function. Moreover, this effect of TRH is reduced or inhibited by atropine, suggesting that cholinergic mechanisms contribute to some of the effects of TRH (ref. 1). It can be shown that general anaesthetics (barbiturates in particular) selectively reduce the sensitivity of cerebral cortical neurones to the excitatory actions of iontophoretically applied acetylcholine[2-4]. This anaesthetic interference with cholinergic, muscarinic excitations has been postulated to participate in the regulation of levels of consciousness[5]. In view of these considerations it might be predicted that TRH would alter the effects of acetylcholine (ACh) on cortical neurones and the experiments described below were designed to test this hypothesis. The results indicate that TRH potentiates the excitatory actions of ACh on cortical neurones.

Female rats (200–210 g, Charles River) were anaesthetised with pentobarbital sodium (50 mg kg^{-1}) with supplemental intraperitoneal injections when required. A small burr hole was drilled over the somatosensory cortex and after reflection of the dura the exposed cortex was covered with 0.5% agar to prevent drying. Neuronal activity was recorded by conventional means through the use of multibarrelled glass micropipettes (tip diameter, 5–10 μm), the centre barrel being filled with 2 M NaCl. The outer barrels of the electrodes were filled immediately before use by centrifugation with solutions of acetylcholine chloride (0.2 M), Na glutamate (0.5 M), carbamyl choline chloride (carbachol, 0.2 M), and TRH (0.05 M in 165 mM NaCl). Automatic current balancing through a barrel containing 2 M NaCl was used to reduce current artefacts.

Spontaneously active neurones, located between 750 μm and 1,250 μm from the cortical surface were examined since the muscarinic nature of ACh excitations of cells in this region of the cortex is well documented[6]. Applications of TRH (ejected with anodal current, 10–70 nA, 30 s–4 min) consistently enhanced the action of iontophoretically applied ACh (10–40 nA, 15–60 s). Thus, on 16 of 18 cells tested, TRH either poten-
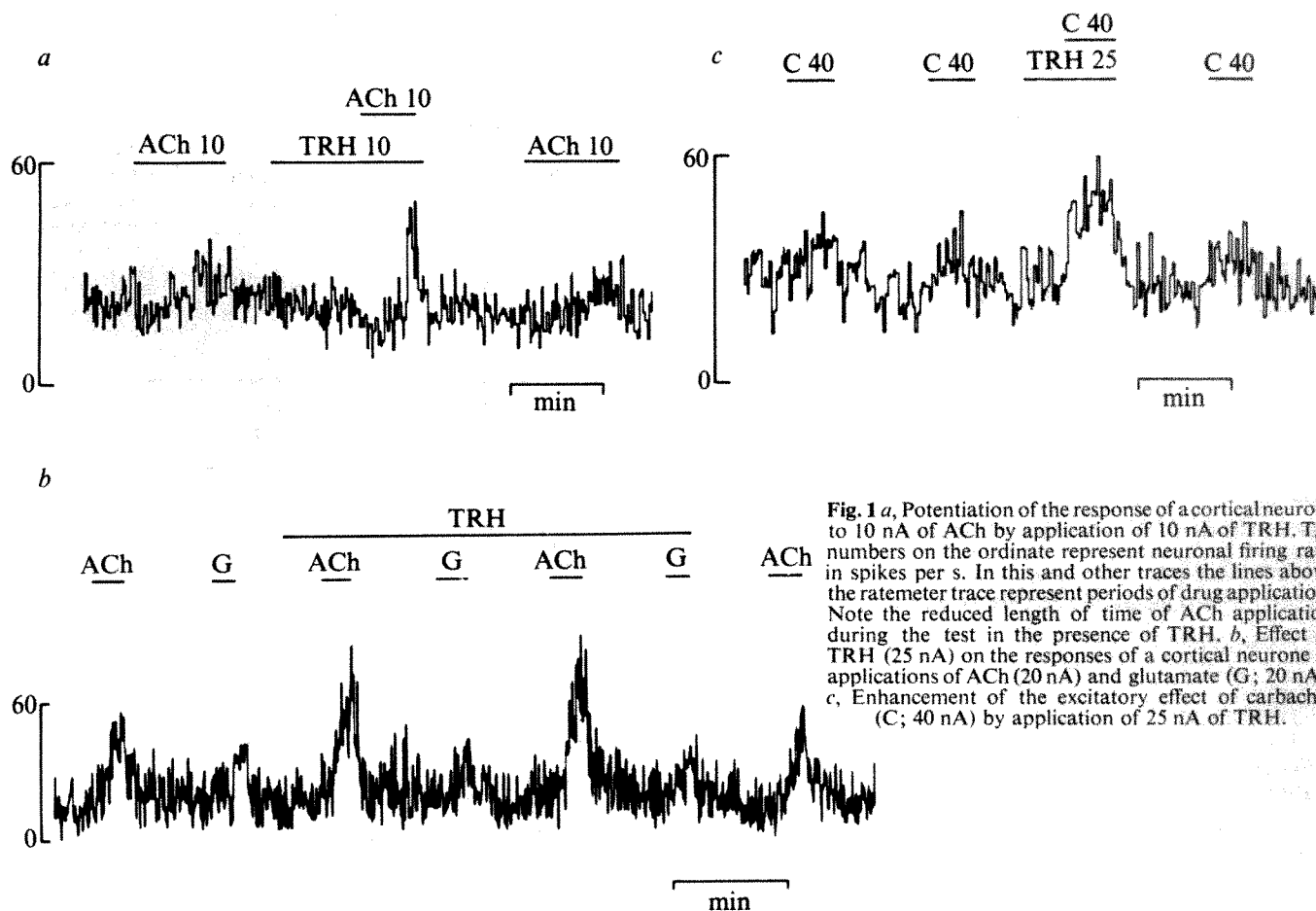
Fig. 1 *a*, Potentiation of the response of a cortical neurone to 10 nA of ACh by application of 10 nA of TRH. The numbers on the ordinate represent neuronal firing rate in spikes per s. In this and other traces the lines above the ratemeter trace represent periods of drug application. Note the reduced length of time of ACh application during the test in the presence of TRH. *b*, Effect of TRH (25 nA) on the responses of a cortical neurone to applications of ACh (20 nA) and glutamate (G; 20 nA). *c*, Enhancement of the excitatory effect of carbachol (C; 40 nA) by application of 25 nA of TRH.

tiated the excitory actions of regularly applied pulses of ACh or converted an inactive ejection current of ACh to a typical ACh excitation. An example of this phenomenon is illustrated in Fig. 1*a*. TRH itself did not directly alter the excitability of the cells tested, nor did it seem to modify substantially the response of 11 cells tested with 20–50 nA of glutamate. Figure 1*b* shows a cell in which the response to ACh was clearly enhanced during the application of TRH while not obviously affecting the response to repeated, submaximal applications of glutamate. Since a possible explanation for the ability of TRH to enhance ACh excitations while not affecting glutamate responses might be that of acetylcholinesterase inhibition, the interactions of TRH and carbachol (not subject to degradation by acetylcholinesterase) were examined. On 11 of 13 cells tested, TRH potentiated the excitatory actions of carbachol (10–40 nA) and a typical example is shown in Fig. 1*c*. This effect of TRH on excitations produced by both ACh and carbachol was consistently brief in onset (sometimes occurring when the TRH and cholinergic agonist were applied simultaneously) and offset.

TRH and other peptides are being increasingly investigated with regard to their possible roles in neural function. The regional and widespread distribution of TRH throughout the neuroaxis[6,7], coupled with the demonstration of a depressant effect of TRH on neurones in a variety of areas[8,9] within the central nervous system (CNS), indicates that TRH might subserve a transmitter function. TRH was not observed to have any direct effects on neuronal excitability of the restricted population of cells sampled in the present study with the ejection currents used, although it clearly and consistently enhanced the excitatory actions of ACh and carbachol. These findings may provide some insight at the neuronal level for the significant anti-anaesthetic effects of TRH in the light of the above considerations, and furthermore, may provide the basis for new approaches in attempts to understand the functional significance of TRH in the CNS. Thus, although not discounting a possible role of TRH as an intercellular mediator, the hormone may subserve a more indirect function to modulate the actions of certain chemical transmitters.

G. G. YARBROUGH

Merck Institute for Therapeutic Research,
West Point, Pennsylvania 19486

1 Breese, G. R., *et al.*, *J. Pharmac. exp. Ther.*, **193**, 11–22 (1975).
2 Krnjevic, K., and Phillis, J. W., *J. Physiol., Lond.*, **166**, 328–350 (1963).
3 Catchlove, R. R. H., Krnjevic, K., and Maretic, H., *Can. J. Physiol. Pharmacol.*, **50**, 1111–1114 (1972).
4 Forrester, P. A., *Br. J. Pharmac.*, **55**(2), 275P–276P (1976).
5 Krnjevic, K., *Physiol. Rev.*, **54**, 418–540 (1974).
6 Winokur, A., and Utiger, R. D., *Science*, **185**, 265–267 (1974).
7 Hökfelt, T., Fuxe, K., Johansson, O., Jeffcoate, S., and White, N., *Eur. J. Pharmac.*, **34**, 389–392 (1975).
8 Dyer, R. G., and Dyball, R. E. J., *Nature*, **486**, 254–255 (1974).
9 Renaud, L. P., and Martin, J. B., *Brain Res.*, **86**, 150–154 (1975).

Incision of ultraviolet-irradiated DNA by extracts of *E. coli* requires three different gene products

IN most organisms pyrimidine dimers induced in DNA by ultraviolet light are removed by excision which is initiated by a repair-specific endonuclease that recognises the damage and makes a strand incision adjacent to the dimer[1–4]. Characterisation of excision-defective mutants of *Escherichia coli* has shown that in this organism early steps of repair are controlled by the *uvrA*, *uvrB* and *uvrC* genes[5]. Although *uvrA* and *uvrB* mutants seem to be incision defective *in vivo*[6,7], it has not been possible to measure any difference in the amount of ultraviolet-endonuclease activity between crude extracts from mutants and wild-type cells[8,9]. After partial purification of wild-type or *uvrC* mutant extracts, however, an ultraviolet endonuclease has

been identified which is absent from *uvrA* and *uvrB* cells[9] (in this communication termed the *uvrAB* endonuclease). The relevance of these results to whole cells is unclear, because recent experiments with permeable cells have shown that *uvrA⁺B⁺*-dependent strand incision requires adenosine-5'-triphosphate (ATP)[10–12], whereas the *uvrAB* endonuclease is independent of ATP[2]. The aim of the present investigation was to observe ATP-dependent ultraviolet-endonuclease activity in a cell-free system. We report here the characterisation in crude extracts of an ATP-dependent ultraviolet-endonuclease activity from *E. coli* and conclude that the activity reflects that the enzyme is essential for repair in whole cells. The activity requires the complementary action of the *uvrA⁺* *uvrB⁺* and *uvrC⁺* products and this has been utilised to establish *in vitro* assays for the individual products of these genes.

In preliminary experiments, extracts made by Brij-58 lysis[13] or lysozyme lysis of freeze-thawed cells[14] were tested for ATP-dependent ultraviolet-endonuclease activity, but results with these extracts were unsatisfactory. Soluble extracts were then prepared by a procedure based on gentle lysozyme lysis of sucrose permeabilised cells (see legend, Fig. 1). This method avoids the use of detergents or freeze-thawing which might affect labile enzyme activities. Typical results obtained by assaying endonuclease activity in extracts from wild-type cells and repair-defective mutants are shown in Fig. 1*a–d*.

In the absence of ATP, extracts from all strains showed similar levels of ultraviolet-endonuclease activity, in agreement with results reported previously[8,9]. Addition of ATP however, caused a fivefold increase in the amount of ultraviolet-endonuclease activity in wild-type extracts, but had only a minor effect on mutant extracts. It is thus clear that the major ultraviolet-endonuclease activity in wild-type extracts requires ATP and depends on functional *uvr* genes. The formation of ultraviolet-induced strand incisions in permeabilised cells has the same characteristics[10,11], and we infer that the activity observed in these extracts reflects the ultraviolet-endonuclease essential for repair in whole cells.

In view of the common deficiency in extracts from cells with mutations mapping at distinct loci, *in vitro* complementation between different mutant extracts was attempted. When extracts from *uvrC* and either *uvrA* or *uvrB* cells were mixed, the ATP-dependent ultraviolet-endonuclease activity was completely recovered (Fig. 1*e–f*). Combined extracts from *uvrA* and *uvrB* cells also displayed activity due to complementation, although usually not to the same extent as combinations including *uvrC* extracts (Fig. 1*g*). Evidently, gene products from *uvrA⁺*, *uvrB⁺* and *uvrC⁺* are required for ATP-dependent

Table 1 Effect of various incubation conditions on the ultraviolet-specific endonucleolytic activity from *uvr⁺* cells

Variation in reaction mixture	Ultraviolet-specific endonuclease activity (%)
Complete reaction mixture	100
−ATP	18
3 mM ATP	115
−KCl	35
3.5 mM Mg²⁺	38
10 mM Mg²⁺	100
20 mM Mg²⁺	75
+25 mM caffeine	10
+10 mM NEM	29

Extracts from cells of strain AB1157 *uvr⁺* were assayed in duplicate for nonspecific and ultraviolet-specific endonucleolytic activity (DNA exposed to 225 J m⁻² of ultraviolet). Net ultraviolet-specific activity was calculated substracting the amount of nicked DNA in non-irradiated samples from the amount of nicked DNA in irradiated samples. The activity obtained in standard assay conditions (Fig. 1) including ATP (1.5 mM) in the reaction mixture was normalised to 100% and the effect of additions (+), omissions (−) or change in concentration of a single component was calculated relative to this value.
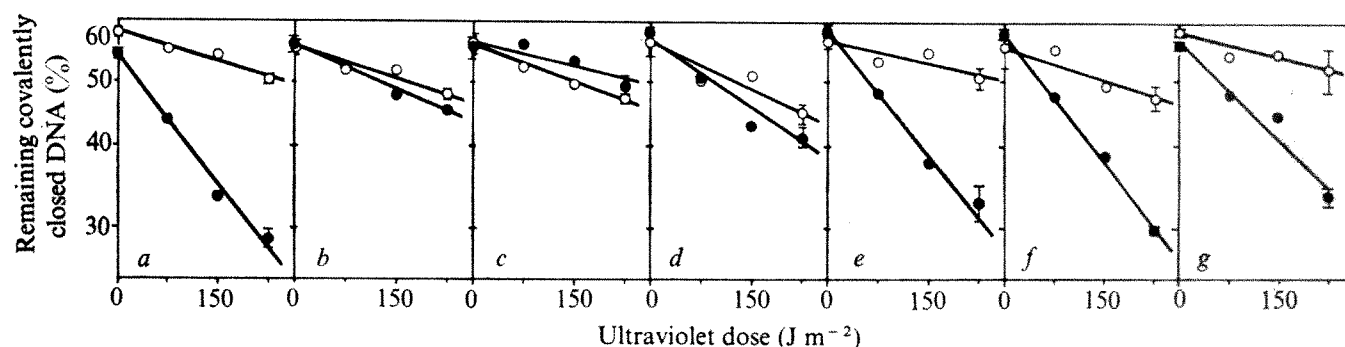
Fig. 1 Ultraviolet-specific ATP-dependent endonucleolytic activity in extracts from excision-proficient and deficient strains of *E. coli* K12. Soluble extracts from strains AB1157 *uvr*⁺ (*a*), AB1886 *uvrA* (*b*), AB1885 *uvrB* (*c*) and AB1884 *uvrC* (*d*), were assayed for endonuclease activity in the absence (○) or presence (●) of ATP (1.5 mM), using the "nicking" assay of Braun *et al.*[9] with the modification that covalently closed circular colE1 DNA (molecular weight = 4.2 × 10⁶) was used as substrate. Plasmid DNA, ³H-labelled, was extracted from growing cells of strain SK2001 *thy*⁻ colE1 by means of a sodium dodecyl sulphate (SDS)–NaCl lysis procedure, and the covalently closed circular form was purified by phenol extraction and banding in CsCl–ethidium bromide. Details of the purification procedure will be described elsewhere (Nissen-Meyer and Seeberg, to be published). For extract preparation cells were first treated with high sucrose as described by Hurwitz *et al.*[19] and then lysed with lysozyme. Approximately 10¹⁰ exponentially growing cells were suspended in 0.25 ml 2.5 M sucrose (0.04 M Tris, pH = 8.0, 10 mM EGTA) and the suspension diluted fivefold in a buffer (50 mM MOPS, pH = 7.5, 100 mM KCl, 1 mM EDTA, 1 mM dithiothreitol), containing 125 μg ml⁻¹ of lysozyme. After 45 min on ice, the lysates were centrifuged at 10,000g for 15 min. The supernatant, subsequently referred to as extract, was completely soluble, essentially DNA free (less than 0.1 % of total DNA), and contained approximately 6 mg of protein. Standard reaction mixtures (130 μl) contained 40 mM MOPS, pH = 7.5, 85 mM KCl, 15 mM MgSO₄, 1 mM EDTA, 0.8 mM dithiothreitol, 5 ng DNA exposed to various doses of ultraviolet light (254 nm), and 10 μl extract. Incubation was for 15 min at 37 °C. Incised DNA was selectively denatured and separated from intact DNA by filtration through nitrocellulose as previously described[9]. The fraction of remaining covalently closed circles was determined by scintillation counting and plotted against ultraviolet exposure to the DNA. *e–g*, Results obtained assaying extract mixtures (1:1) of strains AB1886/AB1884 (*e*), AB1885/AB1884 (*f*) and AB1886/AB1885 (*g*). Each value plotted is the average of two or three independent assays and bars indicate standard deviation.

ultraviolet-endonuclease activity and a mutation at any one of these loci does not prevent the formation of functional products from the other *uvr* genes. The complementation provides an *in vitro* assay for the separate *uvrA*⁺, *uvrB*⁺ and *uvrC*⁺ products. The activity of either product can be assayed by *in vitro* complementation with an extract made from the corresponding mutant, measuring ATP-dependent ultraviolet-endonuclease activity. Purification of the *uvrC*⁺ product is now being attempted using this approach. *In vitro* complementation assays have proved very useful in the purification of proteins active in DNA replication[15,16].

The ATP-dependent ultraviolet-endonuclease observed in crude extracts resembles the partially purified *uvrAB* endonuclease described previously, in that both activities are stimulated by moderate concentrations of salt and require the gene products from *uvrA*⁺ and *uvrB*⁺ (Table 1, ref. 9). The activity in extracts requires ATP, fairly high concentrations of Mg²⁺ and the *uvrC*⁺ product (Table 1, Fig. 1). In contrast, the *uvrAB* endonuclease is not stimulated by ATP, acts in the absence of Mg²⁺ and is present in normal amounts and with normal activity in *uvrC* cells[4,9]. These results can be reconciled if the *uvrC*⁺ product interacts with the *uvrAB* endonuclease to yield ATP-dependent ultraviolet-endonuclease activity. The *uvrC*⁺ product may not have a direct role in the production of single-strand breaks, but may possibly be an ATP-requiring factor which promotes the turnover and thereby amplifies the activity of the endonuclease. The activity of the *uvrAB* endonuclease in the absence of the *uvrC*⁺ product may be detectable in extracts from *uvrC* cells where we repeatedly observed slightly more ultraviolet-endonuclease activity than in *uvrA* or *uvrB* extracts (Fig. 1*d*).

Results with *uvrC* cells *in vivo* have suggested that the *uvrC* function may be involved in a step succeeding incision because single-strand breaks slowly accumulate in the DNA of *uvrC* cells during post-irradiation incubation[7,17]. The rate of formation of these breaks seems to be increased by a mutation affecting ligase activity, suggesting that the *uvrC*⁺ product may prevent ligase reversal of the incision reaction and force the subsequent removal of the dimer[18]. Our results do not eliminate this hypothesis, because ligase may be partially active

in the extracts in spite of the absence of the ligase cofactor nicotinamide adenine dinucleotide in the reaction buffer. It seems clear, however, from the results presented in this communication, that the *uvrC*⁺ product in extracts has an essential role in promoting strand incisions, directly or indirectly. Purification of the different components of the ATP-dependent ultraviolet-endonuclease activity seems necessary before the interrelationship between the *uvrA*⁺, *uvrB*⁺ and *uvrC*⁺ products can be defined in detail.

P.S. acknowledges support from the Royal Society and the MRC.

ERLING SEEBERG
JON NISSEN-MEYER
PETER STRIKE*

Norwegian Defence Research Establishment,
Division for Toxicology,
PO Box 25—N-2007 Kjeller, Norway

*Present address: Department of Genetics, University of Liverpool, Liverpool L69 3BX, UK.

1 Setlow, R. B., and Carrier, W. L., *Proc. natn. Acad. Sci. U.S.A.*, 51, 226–231 (1964).
2 Boyce, R. P., and Howard-Flanders, P., *Proc. natn. Acad. Sci. U.S.A.*, 51, 293–300 (1964).
3 Howard-Flanders, P., *A. Rev. Biochem.*, 37, 175–200 (1968).
4 Grossman, L., Braun, A., Feldberg, R., and Mahler, I., *A. Rev. Biochem.*, 44, 19–43 (1975).
5 Howard-Flanders, P., Boyce, R. P., and Theriot, L., *Genetics*, 53, 1119–1136 (1966).
6 Ogawa, H., Shimada, K., and Tomizawa, J., *Molec. gen. Genet.*, 101, 227–244 (1968).
7 Seeberg, E., and Johansen, I., *Molec. gen. Genet.*, 123, 173–184 (1973).
8 Takagi, Y., *et al.*, *Cold Spring Harb. Symp. quant. Biol.*, 33, 219–277 (1968).
9 Braun, A., and Grossman, L., *Proc. natn. Acad. Sci. U.S.A.*, 71, 1838–1842 (1974).
10 Waldstein, E. A., Sharon, R., and Ben-Ishai, R., *Proc. natn. Acad. Sci. U.S.A.*, 71, 2651–2654 (1974).
11 Seeberg, E., and Strike, P., *J. Bact.*, 125, 787–795 (1976).
12 Strike, P., and Emmerson, P. T., *Molec. gen. Genet.*, 130, 39–45 (1974).
13 Barbour, S. D., and Clark, A. J., *Proc. natn. Acad. Sci. U.S.A.*, 65, 955–961 (1970).
14 Wickner, W., Brutlag, D., Schekman, R., and Kornberg, A., *Proc. natn. Acad. Sci. U.S.A.*, 69, 965–969 (1972).
15 Schekman, R., Weiner, J. H., Weiner, A., and Kornberg, A., *J. biol. Chem.*, 250, 5859–5865 (1975).
16 Morris, C. F., Sinha, N. K., and Alberts, B. M., *Proc. natn. Acad. Sci. U.S.A.*, 72, 4800–4804 (1975).

[17] Kato, T., *J. Bact.*, 112, 1237–1246 (1972).
[18] Seeberg, E., and Rupp, W. D., in *Molecular Mechanisms for the Repair of DNA* (edit. by Hanawalt, P. C., and Setlow, R. B.), 439–441 (Plenum, New York, 1975).
[19] Wickner, R. B., and Hurwitz, J., *Biochem. biophys. Res. Commun.*, 47, 202–211 (1972).

Improved derivative of a phage λ EK2 vector for cloning recombinant DNA

THOMAS *et al.*[1] have constructed a derivative of phage λ (λgt·λC) that is useful for cloning DNA fragments 1,000 to 15,000 base pairs long from other organisms. Their cloning procedure involves cutting the phage DNA into three fragments with *Eco*RI nuclease, removing the central (or "C") fragment, and annealing the outside fragments with an RI digest of foreign DNA. We have recently introduced three amber mutations—*W*am403, *E*am1100, and *S*am100—into λgt·λC with the object of making the phage less likely to encounter a susceptible host in nature[2]. The NIH Advisory Committee on Recombinant DNA Research, on the basis of appropriate tests, has certified λgt*WES*·λC as an EK2 vector[3,4].

In order to extend the usefulness of this vector to the cloning of smaller fragments of DNA (such as synthetic, double-stranded DNA derived from the reverse transcription of mRNA) and to obviate the need for biochemically deleting the 5,400 base pairs central *Eco*RI restriction fragment, λC, we have replaced the λC fragment with the phenotypically inert *Eco*RI fragment λB (refs 5–7).

(In addition to genetic tests necessary to verify the mutations present in λgt*WES*·λC, the right and left restriction fragments of the phage must be purified from the λC fragment before their inclusion in EK2 recombination reactions. The λC fragment encodes genes involved in integration, *int, xis* and *att*, as well as phage-promoted generalised recombination, *red*. This fragment can be accommodated in hybrids with foreign DNA fragments as large as 7,000 base pairs. Thus the λC fragment must be biochemically removed to minimise the opportunity for recombination between bacterial host DNA and a potentially deleterious hybrid as well as to prevent the outgrowth of *red*⁺ reconstituted parent-like recombinant phage. While this purification can be accomplished by sucrose gradient sedimentation or Agarose gel electrophoresis, both procedures are wasteful of time and materials and may damage the single-stranded portion of the RI cleavage site.)

The new derivative retains all the safety features of λgt*WES*·λC but now lacks the λC fragment and the λ elements *int, xis, att* and a portion of the λ *red* gene, which therefore, no longer need to be deleted biochemically. Since the λB fragment contains two immediately adjacent *Sst*I restriction endonuclease sites (S. Goff, personal communication) which enable λgt*WES*·λB to be cleaved into fragments of sufficient length for packaging, it can also be used in conjunction with various "tailing" and "blunt-end ligation" strategies to clone fragments of DNA less than 1,000 base pairs in length. Here we describe the construction and properties of λgt*WES*·λB.

DNA from the two parental phages, λgt*WES*·λC and λgt·λB, were treated with *Eco*RI and permitted to recombine randomly in the presence of T4 DNA ligase (Fig. 1). Plaques obtained after transfection were screened for appropriate markers on three test strains. The desired recombinant contains the left and right arms from the first parent and the λB fragment from the second parent. The presence of amber mutations was recognised by failure to grow on a *sup*O host (*E.c oli* 594), and the presence of *S*am100 by failure to grow on a *sup*E host. The red plaque test was used to detect the absence of the middle λC fragment carrying the *int* and *xis* genes[13]. From candidate recombinants obtained by this screening, the desired phage derivative carrying all three amber mutations was identified by marker rescue studies. Further genetic characterisation has confirmed that this derivative is *red*⁻ and carries *c*Its857 and the *nin*5 deletion (Table 1).

The selected recombinant, λgt *W*am403 *E*am1100 *S*am100·λB, was characterised biochemically by its restriction fragment pattern (Fig. 2). *Eco*RI cleaves λgt*WES*·λB DNA into the three expected fragments: a 21.3 kilobase (kb) left arm, a 13.9 kb right arm and a 4.69 kb middle fragment characteristic of the λB fragment obtained from restriction of λ*c*I857 or λgt·λB DNA. The restriction endonuclease *Sst*I cleaves λ DNA at only two adjacent sites within the λB fragment (S. Goff, personal communication). *Sst*I cleavage of λgt*WES*·λB yields the expected three fragments confirming the presence of the λB fragment. Digestion with *Bam*HI of λgt*WES*·λB and λgt·λB indicates that a single copy of the λB fragment has been inserted in the recombinant in reverse orientation. *Bam*HI cleavage of both phage DNAs yields only four fragments. *Bam*HI would generate a fifth easily-distinguishable band if a second copy of the λB fragment were present, regardless of its orientation with respect to the first. On the basis of their gel mobility, the size of these fragments in λgt·λB are 5.4 kb, 16.7 kb, 6.4 kb, and 11.4 kb as would be expected from the λ *Bam*HI cleavage map[9]. However, in the new recombinant the second and third fragments are 19.3 kb and 3.8 kb, respectively, as would be expected from the fragment orientation shown in Fig. 1.

The new vector, λgt*WES*·λB, greatly simplifies the preparation of hybrid recombinants under EK2 conditions. Since the vector lacks the fragment on which genes for integration and
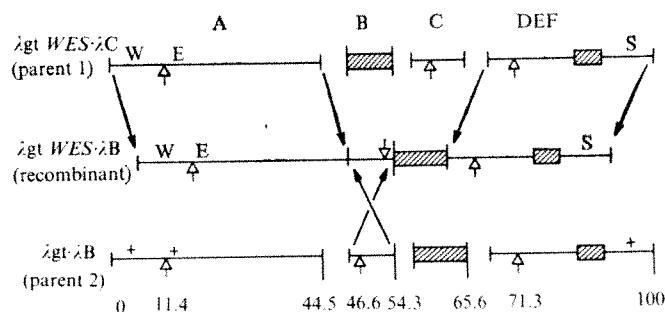
Fig. 1 Construction of λgt *W*am403 *E*am1100 *S*am100·λB *Eco*RI fragments of the two parents λgt*WES*·λC and λgt·λB are represented as separated solid lines and labelled A, B, C, DEF at the top of the figure. *Bam*HI sites in the phage DNA are represented as arrows and both *Eco*RI (ref. 8) and *Bam*HI (ref. 9) target sites are given at the bottom of the figure as percentage λ unit length. Deletions are represented as hatched boxes. The deletion in the right arm of all three phages is *nin*5. The preparation of λgt *W*am403 *E*am1100 *S*am100·λC DNA has been described previously[2]. The construction of λgt·λB has been described[1]. The phage was prepared from plate stocks[10]. λgt·λB DNA was prepared from purified phage as previously described[2]. λgt*WES*·λC and λgt·λB DNA were digested at 50 μg ml⁻¹, and 25 μg ml⁻¹, respectively, by *Eco*RI at 60 U ml⁻¹ in separate 50-μl reactions containing 0.1 M Tris-HCl, pH 7.9; 50 mM NaCl; 12 mM MgCl₂; 0.1 mM EDTA for 30 min at 37 °C. The reactions were stopped by incubation for 5 min at 70 °C. The DNAs digested with *Eco*RI were then mixed in a 50-μl reaction mixture containing 20 μg ml⁻¹ λgt*WES*·λC DNA, 8.3 μg ml⁻¹ λgt·λB DNA, 0.1 M Tris-HCl, pH 7.5; 50 mM NaCl; 12 mM MgCl₂; 0.1 mM EDTA; 10 mM dithiothreitol; 0.1 mM ATP and 50 μg ml⁻¹ bovine serum albumin and recombined by incubation for 48 h at 9 °C with 0.5 U ml⁻¹ T4 DNA ligase (Miles). The reactions were diluted to 1.0 μg ml⁻¹ DNA and used to transfect a T1-resistant isolate of JC5183 (*E. coli* F⁻ *sbc*A5 *rec*B21 *rec*C22 gal⁻ Endo I⁻) (ref. 11) essentially as described in Cameron *et al.*[12]. The transfected cells were plated with the indicator *gal* deletion strain SA825 (*sup*F). Plaques which contained amber mutations were scored by failure to grow on *sup*O strain 594. Plaques containing *S*am100 were scored by poor growth on the *sup*E strain C600. Plaques containing the λB fragment were scored by their formation of white plaques in the test for *int-xis* devised by Enquist and Weisberg[13]. As indicated in Table 1, candidate recombinants were scored for the presence of *E*am1100 by their growth on *gro*E. The presence of all three ambers was confirmed by marker rescue studies described in Table 1.

Table 1 Genetic analysis of λ vectors

	W*	E*·†	(int–xis)‡	red§	cIts857¶	nin5 ‖	S*
λgt·λC	+	+	+	+	+	+	+
λgt·λB	+	+	−	−	+	+	+
λgt Wam403 Eam1100 Sam100·λC	−	−	+	+	+	+	−
λgt Wam403 Eam1100 Sam100·λB	—	—	—	−	+	+	—

The presence of the genetic marker listed was determined by several methods indicated as follows.
*Complementation tests were performed as described by Shimada *et al.*[14]. +, Plaque formation on 594 (*sup*0); −, no plaques observed for the given complementation test.
†Phage carrying the *Eam*1100 mutation will form plaques (+) on strain Ymel *groE* whereas phage lacking the mutation will not (−) (refs. 2 and 15).
‡The red plaque test for λ *int* and *xis* function was used as described by Enquist and Weisberg[13]. +, Formation of red plaques, thus (*int–xis*)⁺; −, white plaques or (*int–xis*)⁻.
§Phage lacking the λ *red* function will not form plaques on *Feb*⁻ hosts[16]. We used Ymel *lig*ts7 at 32 °C as the *Feb*⁻ host. +, Phage growth on this host; −, no growth.
¶Phages carrying the *c*Its857 mutation form turbid plaques at 32 °C and clear plaques at 42 °C. +, Phages with this phenotype.
‖The presence of the *nin*5 deletion was determined as described in Enquist *et al.*[2]. +, Phages carrying the *nin*5 deletion.

generalised recombination reside, the larger λ fragments need not be purified and the possibility of low-level λC fragment contamination is eliminated. As with λgt*WES*·λC, this vector should accommodate DNA fragments as large as 13–15 kb.

The susceptibility of λgt*WES*·λB DNA to *Sst*I nuclease has
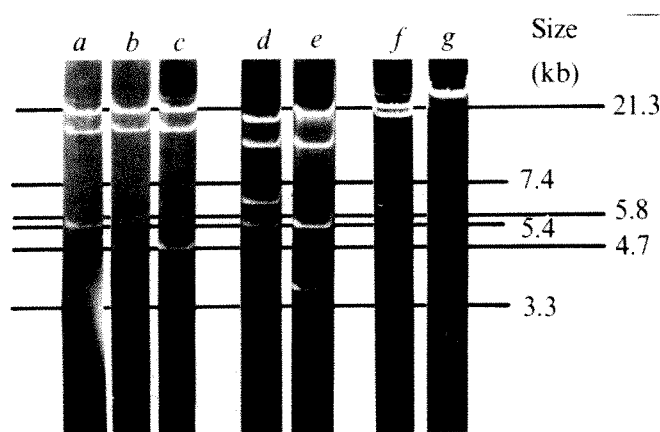
Fig. 2 Restriction enzyme analysis of λgt *W*am403 *Eam*1100 *Sam*100·λB DNA. *a*, λgt *W*am403 *Eam*1100 *Sam*100·λC (*Eco*RI); *b*, λgt *W*am403 *Eam*1100 *Sam*100·λB (*Eco*RI); *c*, λgt·λB (*Eco*RI); *d*, λgt·λB (*Bam*HI); *e*, λgt *W*am403 *Eam*1100 *Sam*100·λB (*Bam*HI); *f*, λgt *W*am403 *Eam*1100 *Sam*100·λB (*Sst*I); *g*, λgt *W*am403 *Eam*1100 *Sam*100·λC (*Sst*I). The mobility of standard λ*c*I857 *Eco*RI fragments in this gel system is represented by the horizontal lines and then sizes are given in kilobase pairs. λgt *W*am403 *Eam*1100 *Sam*100·λB phage was prepared by liquid infection initiated at a multiplicity of 0.01 and a cell density of 2 × 10⁸ ml⁻¹. After an 8-h incubation at 37 °C, the suspension was treated with 1% CHCl₃ for 10 min at 37 °C, brought to 10 mM MgCl₂, and centrifuged for 10 min at 20,000*g* to remove cell debris. λ*c*I857, λgt *W*am403 *Eam*1100 *Sam*100·λC and λgt·λB phage were prepared from induced lysogens or plate stocks[10]. Phage purification and DNA isolation have been previously described[2]. DNA solutions (30 to 830 µg ml⁻¹) are stored at 0–5 °C in 10 mM Tris-HCl, pH 7.5; 0.1 mM EDTA. *Eco*RI was purified from *E. coli* BRY 13.3/1100.5 (to be described elsewhere). Purified *Sst*I was the gift of Dr Steven Goff. *Bam*HI was obtained from Bethesda Research Labs. The standard restriction reaction contained 0.3 µg DNA in 25 µl and sufficient enzyme to cleave 1.0 µg DNA. The *Eco*RI buffer was 0.1 M Tris-HCl, pH 7.9; 50 mM NaCl; 0.1 mM EDTA; 12 mM MgCl₂. The *Sst*I and *Bam*HI buffer was 20 mM Tris-HCl, pH 7.5; 7.0 mM MgCl₂; 60 mM NaCl; 2 mM β-mercaptoethanol. The reactions were incubated for 45 min at 37 °C. After subsequent incubation for 5 min at 75 °C, the reactions were electrophoresed through 1.0% Agarose (SeaKem) slab gels as described previously[17]. The DNA was stained by immersion in 0.0025% ethidium bromide for 3 min and made visible under an ultraviolet lamp. The small 1.0-kb band generated by the *Sst*I cleavage is not readily visible in the contact print of this gel system.

two implications. First, it should be possible to treat the vector with *Sst*I as well as *Eco*RI before hybrid formation to eliminate the bulk of parental recombinants. The *Eco*RI λB fragment when reduced to two fragments each with one *Eco*RI and one *Sst*I end will still anneal to other *Eco*RI-ended fragments but will require the ligation of two *Eco*RI cleavage sites and an *Sst*I cleavage site in order to form a viable transfectant. The introduction of foreign DNA fragments containing only *Eco*RI ends requires only the annealing and ligation of two *Eco*RI sites. Second, the two arms created by *Sst*I cleavage of λgt*WES*·λB DNA are sufficiently longer than the *Eco*RI arms to allow them to be resealed into plaque-forming DNA. *Sst*I arms thus represent a potential EK2 vector for insertion of low molecular weight double-stranded DNA molecules which are too short for insertion into the *Eco*RI arms alone.

We thank Dr R. W. Davis for λgt·λB phage, Dr S. Goff for purified *Sst*I, Dr J. Davies for *E. coli* BRY 13.3/1100.5, Dr A. J. Clark for *E. coli* strain JC5183, and Drs R. Weisberg and N. Sternberg for helpful discussion. We thank Ms Catherine Kunkle for her assistance in preparing this manuscript.

DAVID TIEMEIER
LYNN ENQUIST
PHILIP LEDER

Laboratory of Molecular Genetics,
National Institute of Child Health
and Human Development,
National Institutes of Health,
Bethesda, Maryland 20014

1 Thomas, M., Cameron, J. R., and Davis, R. W., *Proc. natn. Acad. Sci. U.S.A.*, 71, 4579–4583 (1974).
2 Enquist, L., Tiemeier, D., Leder, P., Weisberg, R., and Sternberg, N., *Nature*, 259, 596–598 (1976).
3 *Annex to Nucleic Acid Recombinant Scientific Memoranda*, September 15, 1975–January 15, 1976.
4 *Nucleic Acid Recombinant Scientific Memoranda*, May 15–September 15, 1976 (in the press).
5 Parkinson, J. S., *J. molec. Biol.*, 56, 385–401 (1971).
6 Davidson, N., and Szybalski, W., *The Bacteriophage Lambda* (edit. by Hershey, A. D.), 45–82 (Cold Spring Harbor Laboratory, New York, 1971).
7 Skalka, A., *The Bacteriophage Lambda* (edit. by Hershey, A. D.), 535–547 (Cold Spring Harbor Laboratory, New York, 1971).
8 Thomas, M., and Davis, R. W., *J. molec. Biol.*, 91, 315–328 (1975).
9 Haggerty, D. M., and Schleif, R. F., *J. Virol.*, 18, 659–663 (1976).
10 Schrenk, W. J., and Weisberg, R. A., *Molec. gen. Genet.*, 137, 101–107 (1975).
11 Barbour, S. D., Nagaishi, H., Templin, A., and Clark, A. J., *Proc. natn. Acad. Sci. U.S.A.*, 67, 128–135 (1970).
12 Cameron, J. R., Panasenko, S. M., Lehman, I. R., and Davis, R. W., *Proc. natn. Acad. Sci. U.S.A.*, 72, 3416–3420 (1975).
13 Enquist, L., and Weisberg, R., *Virology*, 72, 147–153 (1976).
14 Shimada, K., Weisberg, R. A., and Gottesman, M. E., *J. molec. Biol.*, 63, 483–503 (1972).
15 Sternberg, N. J., *J. molec. Biol.*, 76, 1–23 (1973).
16 Zissler, J., Signer, E., and Schaefer, F., *The Bacteriophage Lambda* (edit. by Hershey, A. D.), 455–468 (Cold Spring Harbor Laboratory, New York, 1971).
17 Helling, R. B., Goodman, H. M., and Boyer, H. W., *J. Virol.*, 14, 1235–1244 (1974).

RNA pattern of 'swine' influenza virus isolated from man is similar to those of other swine influenza viruses

In February 1976, during an epidemic of influenza among military recruits at Fort Dix, New Jersey, five virus isolates were obtained which contained haemagglutinin and neuraminidase antigens indistinguishable from those of isolates of swine influenza virus[1]. Subsequent evidence of elevated serum antibody titre against 'swine' influenza virus in hundreds of individuals at Fort Dix[2] prompted the US government to launch a national immunisation programme against 'swine' influenza virus. One current theory suggests that new, pandemic strains of influenza virus result from genetic recombination of human influenza A viruses and strains of virus that are resident in animal populations[3,4]. According to this hypothesis, recombinants may emerge containing genes derived from the animal virus which code for novel surface proteins, and genes derived from the human virus which confer properties of virulence in man. We have demonstrated that the genetic composition of recombinant viruses could be determined by comparison of RNA patterns of recombinant and parent viruses on urea-polyacrylamide gels[5-7]. We have therefore used the same techniques to explore the possibility that the 'swine' virus isolated from recruits at Fort Dix is a recombinant of swine influenza virus from pigs and H3N2 influenza A virus currently prevalent in man.

Figure 1 shows the RNA patterns of three low passage isolates of 'swine' virus obtained from three patients at Fort Dix. All three isolates are identical with respect to the migration of the eight corresponding RNA segments on urea-polyacrylamide gels, confirming previous observations regarding the reproducibility of the RNA patterns in this system.

We compared the RNA pattern of the New Jersey 'swine' virus isolate with those of four influenza viruses isolated from swine and with that of a H3N2 virus isolated from recruits during the same epidemic at Fort Dix in February, 1976. Figure 2 demonstrates the results of one experiment, but it should be noted that at least three different ³²P-labelled RNA preparations were made from each virus and that numerous gels were run to verify the results.

Figure 2 shows that the RNA pattern of the New Jersey 'swine' virus, A/NJ/11/76, isolated from man (lane 2) is virtually identical to that of the swine virus isolate A/swine/Taiwan/1/75 (lane 3), and except for the fifth RNA segment closely resembles the pattern produced by another swine virus isolate, A/swine/Tennessee/1/75 (lane 4). Comparison of a third swine virus, A/swine/Wisconsin/1/73 (lane 7), and the New Jersey 'swine' virus (lane 8) reveals close similarity of the top four RNA segments as well as slight differences in mobility of RNA segments five to eight. In contrast, each of the RNA segments of the H3N2 virus, A/NJ/743/76 isolated from recruits during the same epidemic at Fort Dix (lane 1) could be distinguished from the corresponding RNA segments of the New Jersey 'swine' virus (lane 2), except possibly for RNA segment three. Note that the fourth segment of both viruses codes for haemagglutinin and the fifth segment of A/NJ/743/76 (H3N2) and the sixth of A/NJ/11/76 (HswlN1) code for the respective neuraminidases[7,8]. In particular, swine viruses were characterised by a broad diffuse band at the top in which RNAs 1 and 2 cannot easily be resolved. Furthermore, the eighth RNA segments of all swine viruses migrate faster than the corresponding segments of H3N2 viruses.

Figure 2 also demonstrates the RNA pattern of an H3N2 virus isolated from swine in 1970, A/swine/Taiwan/1/70 (lane 6), which except for RNA segment seven is clearly distinguishable from patterns produced by the other swine virus isolates. This observation is in accord with the previous assumption that this strain, A/swine/Taiwan/1/70, is a virus of man which crossed

species barriers and infected pigs[10]. (Comparison of the RNAs of this virus with those of A/Hong Kong/8/68 virus on other gels reveal very similar patterns.)

Figure 3 is a photograph of another gel demonstrating the RNA patterns of some of the viruses not completely resolved
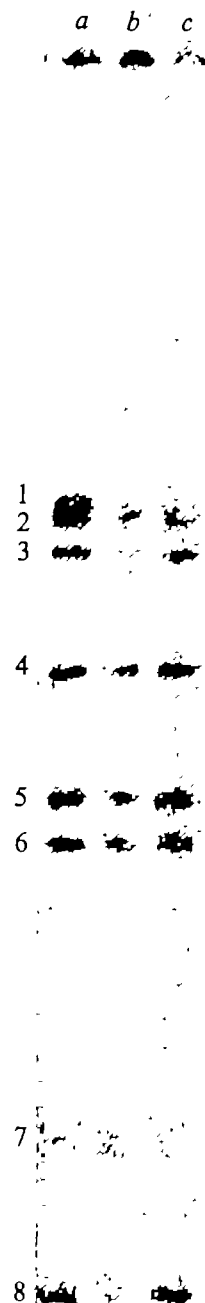
Fig. 1 Analysis of RNAs of three New Jersey 'swine' influenza virus isolates on a 6 M urea–polyacrylamide gel. RNAs were labelled with ³²P during a one-step growth cycle in MDCK (canine kidney) cells as described previously[4-7]. Electrophoresis on a 6 M urea (2.6%)–polyacrylamide gel was carried out according to previously published procedures[5-7]. Migration was from top to bottom. The RNA segments are numbered 1 to 8. *a*, RNA of influenza A/NJ/11/76 (New Jersey 'swine'; HswlN1) virus; *b*, RNA of influenza A/NJ/8/76 (New Jersey 'swine'; HswlN1) virus; *c*, RNA of influenza A/NJ/9/76 (New Jersey 'swine'; HswlN1) virus. Throughout the manuscript we are using the system of nomenclature recommended by the WHO[9]. According to this system only viruses isolated from swine carry the strain designation swine. The strains isolated from recruits at Fort Dix, which are characterised by swine virus surface antigens, are, however, referred to as New Jersey 'swine' viruses in this report.
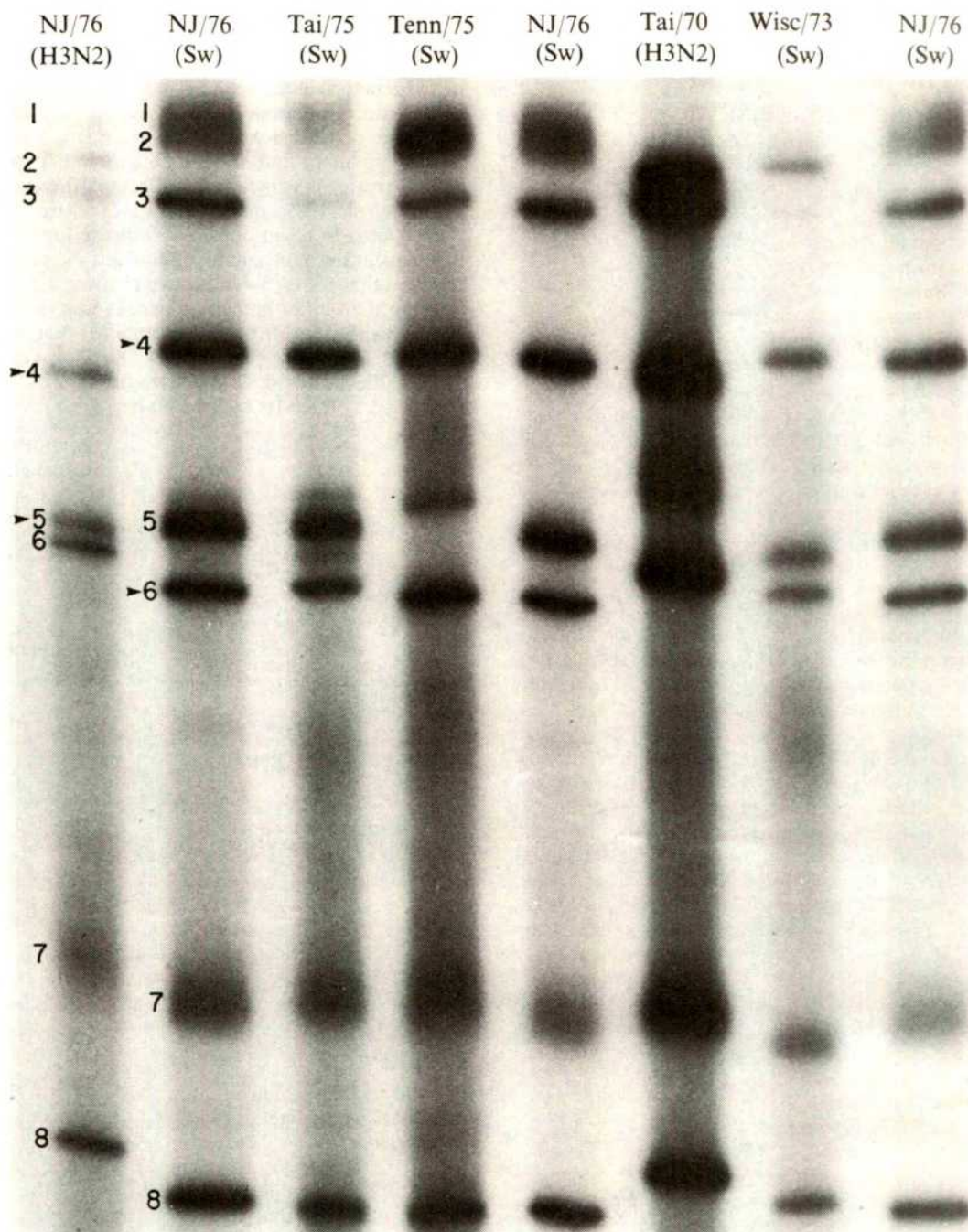
Fig. 2 Comparison of the RNA pattern of the New Jersey 'swine' influenza virus with those of A/NJ/743/76 (H3N2) virus and four viruses isolated from pigs. Conditions of electrophoresis on the 6-M urea–polyacrylamide gel are as in Fig. 1. The RNA segments of A/NJ/743/76 and the New Jersey 'swine' virus (A/NJ/11/76) are numbered 1 to 8. The haemagglutinin genes (RNA 4) and the neuraminidase genes (RNA 5 or 6) of these two viruses are identified by arrows[7,8]. Lanes: 1, RNA of influenza A/NJ/743/76 (H3N2) virus; 2, 5 and 8, RNA of A/NJ/11/76 (New Jersey 'swine'; Hsw1N1) virus; 3, RNA of influenza A/swine/Taiwan/1/75 (Hsw1N1) virus; 4, RNA of influenza A/swine/Tennessee/1/75 (Hsw1N1) virus; 6, RNA of influenza A/swine/Taiwan/1/70 (H3N2) virus; 7, RNA of influenza A/swine/Wisconsin/1/73 (Hsw1N1) virus; note that in the conditions of labelling and growth used, virus with equimolar distribution of RNAs is not always obtained.

in Fig. 2. (We found it difficult to obtain clear resolution of each sample when many viruses are compared on a single gel.) Again it can be seen that the pattern of the New Jersey 'swine' virus in lane 2 resembles that of A/swine/Wisconsin/1/73 (lane 3) and can be distinguished from that produced by either the 1976 H3N2 strain, A/NJ/743/76, (lane 1) or the H3N2 strain, A/swine/Taiwan/1/70, isolated from swine in 1970 (lane 4).

Based on this preliminary analysis, we suggest that the New Jersey 'swine' strain is closely related to other swine viruses—particularly the A/swine/Taiwan/1/75 virus—with respect to all of its 8 genes and most likely was not derived by recombination involving the current human H3N2 strain.

We recognise potential pitfalls in this kind of RNA analysis. First, slight differences in migration rates of comparable RNA segments of two viruses may reflect only minor differences in nucleotide sequences, masking genetic homology. Second, it is possible that the New Jersey 'swine' virus isolate is a recombinant which derives some of its genes from an unrecognised H3N2 parent. Although we have compared the RNAs of five different H3N2 strains with that of the New Jersey 'swine' virus we are aware that these represent only a small sample of potential parents. In addition, we cannot exclude the fact that some small differences in RNAs may not be detectable by this method. Consequently a definitive analysis of the genetic composition
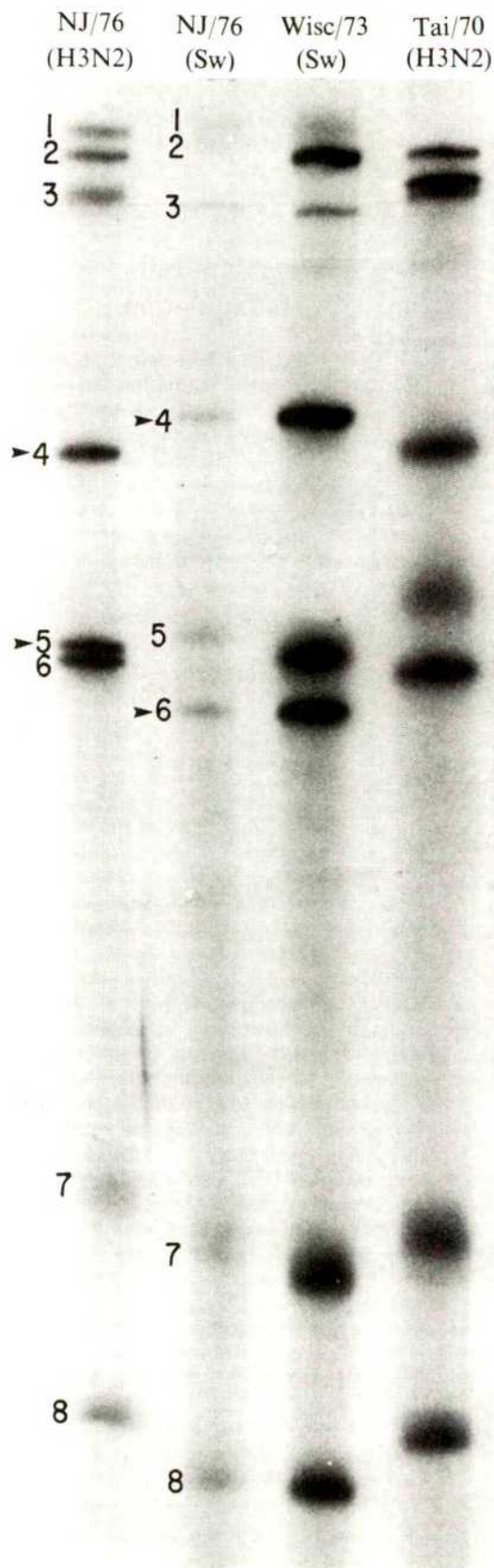
Fig. 3 Analysis of the RNAs of the New Jersey 'swine' influenza virus with those of A/NJ/743/76 (H3N2) virus and two isolates from pigs. Conditions of electrophoresis are as in Fig. 2 but new RNA preparations were used. The RNA segments of A/NJ/743/76 and the New Jersey 'swine' virus (A/NJ/11/76) are numbered 1 to 8. Arrows identify the haemagglutinin genes (RNA 4) and the neuraminidase genes (RNA 5 or 6)[7,8]. Lanes: 1, RNA of influenza A/NJ/743/76 (H3N2) virus; 2, RNA of influenza A/NJ/11/76 (New Jersey 'swine'; Hsw1N1) virus; 3, RNA of influenza A/swine/Wisconsin/1/73 (Hsw1N1) virus; 4, RNA of influenza A/swine/Taiwan/1/70 (Hsw1N1) virus.

of the New Jersey 'swine' virus can only be obtained by more sensitive and elaborate techniques such as oligonucleotide fingerprinting of all of the RNAs or peptide mapping of each of the gene products.

It is conceivable that the New Jersey 'swine' virus acquired properties of virulence in man after selection during sequential passage in man; but if recombination between human and animal strains is generally required for the emergence of new strains virulent for man, we would regard the New Jersey 'swine' virus to be an unlikely candidate for the next influenza pandemic. (Since the submission of this manuscript it has been reported that the New Jersey 'swine' virus is in fact less virulent for man than other human influenza virus isolates[11].)

We thank Ms Marlene Lin and Ms Kaye Leitzinger for technical assistance, and Drs W. R. Dowdle, B. Easterday, M. Goldfield and R. G. Webster for viruses. This work was supported by the NIH and the NSF.

PETER PALESE
JEROME L. SCHULMAN

Department of Microbiology,
Mount Sinai School of Medicine of CUNY,
New York New York, 10029

1 *Morbidity and Mortality Weekly Report*, US Department of Health Education and Welfare, **25**, 47–48 (1976).
2 *Morbidity and Mortality Weekly Report*, US Department of Health Education and Welfare, **25**, 63–64 (1976).
3 Kilbourne, E. D., *Science*, **160**, 74–76 (1968).
4 Laver, W. G., and Webster, R. G., *Virology*, **51**, 383–391 (1973).
5 Palese, P., and Schulman, J. L., *J. Virol.*, **17**, 876–894 (1976).
6 Schulman, J. L., and Palese, P., *J. Virol.* (in the press).
7 Palese, P., and Schulman, J. L., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 2142–2146 (1976).
8 Palese, P., Ritchey, M. B., Schulman, J. L., and Kilbourne, E. D., *Science* (in the press).
9 *Bull. WHO*, **45**, 119–124 (1971).
10 Kundin, W. D., *Nature*, **228**, 857–857 (1970).
11 Beare, A. S., and Craig, J. W., *Lancet*, ii, 4–5 (1976).

Errata

In the article "Three step yielding of load clamped mammalian cardiac muscle" by P. Housmans and D. L. Brutsaert (*Nature*, **262**, 56; 1976) line 30 on page 57, . . . length change, of phase 2 and the magnitude and speed of . . . should be replaced by . . . the clamp (Fig. 2, lower), disclosed that the length change of . . .

In the article "The zoo: 150 years of research" by L. Harrison Matthews (*Nature*, **261**, 281; 1976) the illustration at the top of page 282 is of the walrus *Trichecus rosmarus* and not the bearded seal as indicated.

matters arising

Bicuculline and visual responses

DUFFY *et al.*[1] proposed that the influence of intravenously administered bicuculline on the receptive field properties of neurones in the visual cortices of cats which have experienced monocular deprivation indicates the "active inhibition of the relatively intact input from the amblyopic eye" by information arising from the normal eye.

In investigations of this type, it hardly seems necessary to expose unanaesthetised, paralysed animals (admittedly locally anaesthetised "to avoid possible confounding" of the results) to bicuculline administered systemically, particularly when its action "was often complicated by its potent convulsive effects" The microelectrophoretic administration of bicuculline or other GABA antagonists near physiologically identified neurones in various regions of the visual system would be likely to provide more definitive evidence regarding the involvement and localisation of GABA-mediated inhibitory mechanisms in binocular vision, and their perturbation as a consequence of monocular deprivation, with considerably less distress to the experimental animal.

DAVID R. CURTIS

Department of Pharmacology,
Australian National University,
Canberra, Australia

1 Duffy, F H , Snodgrass, S R., Burchfiel, J L , and Conway, J. L , *Nature*, 260, 256–257 (1976)

DUFFY ET AL REPLY—Curtis has raised[1] at least two issues. He seems to imply that the use of locally anaesthetised animals is improper or unwarranted and, second, he raises the issue of the proper place of iontophoresis in pharmacological studies of the amblyopic animal.

We feel that experimenters have an obligation to minimise an animal's discomfort. No procedure, however, can successfully claim total freedom from trauma. General anaesthesia is probably the best in this regard and should be used whenever possible. General anaesthetics do act on the brain and, in our experience, modify the receptive field characteristics in the visual and somatosensory system. Furthermore,

anaesthetic agents may interact with other experimentally administered drugs, thereby confounding experimental findings and possibly requiring additional experimentation. Since most neurophysiological and neuropharmacological investigations result in an animal's death, we feel that experimenters have an equal obligation to minimise the number of animals used.

In the course of behavioural studies of amblyopic cats, we have administered bicuculline to a number of awake and unrestrained animals. Bicuculline seems to have a sedative effect at low and medium dosage levels. At higher levels, epileptic activity in the EEG and convulsions occur quite suddenly and result in an immediate loss of consciousness Cats do not seem unduly distressed at any time For all these considerations, we felt that the locally anaesthetised preparation was best for our initial experimentation.

We agree with Curtis that iontophoresis is especially useful for discerning the location of a given drug effect within the nervous system, but feel that it has some problems as an exploratory technique. For example, a negative iontophoretic study with bicuculline in amblyopic cats would not have great meaning unless many regions were sampled which would necessitate the use of more animals. We would also point out that when the data obtained with iontophoretic and intravenous drug administration are in apparent conflict, it is by no means certain that the iontophoretic data give a correct perspective. For example, we cite the recent work of Ben-Ari and Kelly[2] where iontophoretic, but not intravenous, flupenthixol blocked the response to iontophoretically applied dopamine. We do plan to make use of iontophoresis in the future and expect that it will be useful in elucidating the detailed mechanism of effect.

Finally we agree that intravenous bicuculline is a toxic agent and have been seeking a safer agent with a longer duration of action. We believe that intravenous administration of ammonium salts may meet these requirements.

Children's Hospital Medical Center,
Boston, Massachusetts

1 Curtis, D. R , *Nature*, 263, 531 (1976)
2 Ben-Ari, X , and Kelly, X., *J. Physiol* , *Lond* , 256, 1 (1976)

Mammalian cell growth regulation

HOLLEY has proposed that mammalian cell growth regulation in culture is regulated primarily by the depletion of diffusable resources, especially polypeptide factors; that the responsiveness of cells to such factors is density dependent; and that this responsiveness changes in a characteristic manner following transformation.

If this theory is correct, the following must be true: (1) that normal and malignant cells differ demonstrably in their growth regulatory policies; (2) that growth control involves a conventional density dependence; (3) that growth inhibition at high density is not merely the result of culture starvation by careless investigators; and (4) that mechanisms other than diffusable substances contribute little to growth regulation.

It may fairly be stated that there is no *in vitro* cell characteristic which has ever been rigorously demonstrated to correlate with and be diagnostic of the malignant state for a broad spectrum of cell types. The following *in vitro* indices are not broad spectrum malignancy correlates: saturation density, multilayering, contact growth inhibition, growth rate, and growth in low serum[1-10]; cell adhesiveness and surface charge[11-14]; lectin agglutinability[15-18]; fibrinolytic activity[19,20]; cell morphology and karyotype[8,9,21,22]. A recalculation of chi-square values for the data of ref. 23, assuming 3T3 cells to be malignant because of their production of tumour angiogenesis factor[24] and *in vivo* tumorigenicity[8], casts grave doubt on the lack of anchorage dependence as a valid tumorigenicity index. Similarly, the status of the LETS-SF complex is in doubt because many studies of it have used as "normal" controls either cell lines of untested tumorigenicity or lines known (3T3 (ref. 2), BHK21 (refs 9 and 25) or suspected (Wi38 ref. 24)) to be tumorigenic. Holley's argument about growth policy differences between normal and malignant cells is clearly without foundation.

Holley contends that density-dependent growth regulation is "due to a quantitative increase in the requirements for macromolecular growth factors as cell density increases." His only basis for this statement is the citation of three references which speculate about growth regulation by the density-dependent

control of cell surface area and nutrient uptake, a theory which now seems unlikely[34,35], and "wound healing" in monolayer culture. This wound healing, once thought to be density dependent, now seems to involve the destruction of an external diffusion barrier near the cell surface[36]. In certain culture systems, growth inhibition definitely does not involve density dependence[27].

Holley defines density-dependent growth regulation as a tendency for cells "to grow to a 'saturation' density and then stop growing." This definition is unsatisfactory. Density dependence is a change per cell in the value of a cellular property as population density changes. As a rule, a density dependence of growth operates at all population densities and may involve either or both the promotion and inhibition of growth rate[10,26,27]. In culture, the inhibition of cell growth usually occurs even when nutrients and diffusable factors are in excess and not rate limiting[3-5,10,27-33]. Although in well fed cultures growth inhibition is occasionally observed[4], it occurs at much higher densities than are normally employed and probably reflects the restriction of diffusional transport either by an external diffusion barrier[36] or by multilayering. Growth inhibition and its density dependence in culture are not usually the result of resource depletion except for the special situation in which the investigator fails to provide adequate nutrient to his cultures. In general, both 'saturation' density and quiescence are starvation artefacts[3-6,10,27-33].

Holley's acceptance of the physiological significance of certain diffusable growth affectors is uncritical. Of the materials he cites, only the nerve and epidermal growth factors (NGF, EGF) are clearly growth regulators *in vivo*. Plant lectins certainly are not, cyclic AMP probably is not[37-40], while the role of fibroblast growth factor, insulin, hydrocortisone, prostaglandins, antigens, and proteolytic enzymes *in vivo* is unknown. Since blood vessels are selectively permeable, the mere presence of a growth affector in serum does not mean that the substance regulates cell growth. Its presence in the interstitial fluid surrounding a suspected target cell must also be demonstrated. Similarly, a growth affector from one species of vertebrate cannot be considered physiological when applied to cells from another, as is the case with several of the studies Holley cites. Even Holley's use of the term 'polypeptide factors' is in error. Such factors (S₂, erythropoietin, EGF, NGF, S₁) are often multimers and often of very high molecular weight (26, 46, 74, 140, 600 kdalton respectively); or very small and not polypeptides at all (putrescine, uridine, adenine).

Finally, recent evidence raises the possibility that the extracellular matrix

---

Matters arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words and the briefest of replies, to the effect that a point is taken, should be considered.

---

may contribute to growth regulation in some systems[41-43], and that contact interactions may also[27], though not in the manner postulated by Todaro and others.

In summary, growth regulation is complex, confusing, and not at all understood. Certain diffusable substances (NGF, EGF, erythropoietin, colony-stimulating activity) are clearly physiological growth regulators. And certainly density-dependent growth regulation occurs in many culture systems, though it is far more complex and often differs fundamentally from Holley's description of it. While density-dependent growth regulation is sometimes mediated by diffusable substances, it is probably at times mediated by matrix and contact interactions as well. There is at present no evidence for density-dependent increases by cells in their resource requirements, nor is there evidence that resource depletion regulates growth *in vivo*. Even in culture, resource depletion seldom contributes to growth regulation of well fed cells. Finally, no general difference between normal and malignant cell growth regulation in culture has ever been rigorously proved.

PHILIP SKEHAN

*Department of Pharmacology,
University of Colorado Medical Centre,
Denver, Colorado 80220*

1 Holley, R. W., *Nature*, **258**, 487 (1975).
2 Boone, C. W., Takeichi, N., Paranjpe, M., and Gilden, R., *Cancer Res.*, **36**, 1626 (1976).
3 Kruse, P. F. Jr, Myhr, B. C., Johnson, J. E., and White, P. B., *J. natn. Cancer Inst.*, **31**, 109 (1963).
4 Kruse, P. F. Jr and Miedema, E., *J. Cell Biol.*, **27**, 273 (1965).
5 Kruse, P. F. Jr, Whittle, W., and Miedema, E., *J. Cell Biol.*, **42**, 113 (1969).
6 Holley, R. W., and Kiernan, J. A., *Proc. natn. Acad. Sci. U.S.A.*, **60**, 300 (1968).
7 Sanford, K. K., Dunn, T. B., Covalesky, A. B., Dupree, L. T., and Earle, W. R., *J. natn. Cancer Inst.*, **26**, 331 (1961).
8 Sanford, K. K., *Int. Rev. Cytol.*, **18**, 249 (1965).
9 Defendi, V., Lehman, J., and Kraemer, P., *Virology*, **19**, 592 (1963).
10 Skehan, P., *Expl Cell Res.*, **86**, 237 (1974).
11 Weiss, L., *Proc. 7th Canadian Cancer Conf.*, 292, (Pergamon, Toronto, 1967).
12 Weiss, L., *The Cell Periphery, Metastasis, and Other Contact Phenomena* (North-Holland, Amsterdam, 1967).
13 Vassar, P. S., *Lab. Invest.*, **12**, 1072 (1963).
14 Easty, G. C., in Amrose, E. J., and Roe, F. J. C., *Biology of Cancer*, ch. 4 (Wiley, New York, 1975).

15 Sivak, A., and Wolman, S. R., *In Vitro*, **8**, 1 (1972).
16 Glimelius, B., Westermark, B., and Ponten, J., *Int. J. Cancer*, **14**, 314 (1974).
17 Glimelius, B., Nilsson, K., and Ponten, J., *Int. J. Cancer*, **15**, 888 (1975).
18 Berman, L. D., *Int. J. Cancer*, **15**, 973 (1975).
19 Goldberg, A. R., *Cell*, **2**, 95 (1974).
20 Unkeless, J. C., Gordon, S., and Reich, E., *J. exp. Med.*, **139**, 834 (1974).
21 Sanford, K. K., *et al.*, *J. natn. Cancer Inst.*, **11**, 35 (1950).
22 Evans, C. E., and DiPaolo, J. A., *Cancer Res.*, **35**, 1035 (1975).
23 Freedman, V. H., and Shin, S., *Cell*, **3**, 335 (1974).
24 Klangsbrun, M., Knighton, D., and Folkman, J., *Cancer Res.*, **36**, 110 (1976).
25 MacPherson, I., *J. natn. Cancer Inst.*, **30**, 795 (1963).
26 Rein, A., and Rubin, H., *Expl Cell Res.*, **49**, 666 (1968).
27 Skehan, P., *Expl Cell Res.*, **97**, 184 (1976).
28 Westermark, B., *Expl Cell Res.*, **69**, 251 (1971).
29 Zinninger, G. G., and Little, J. B., *Cancer Res.*, **33**, 2343 (1973).
30 Castor, L. N., *J. Cell Physiol.*, **72**, 161 (1968).
31 Castor, L. N., *J. Cell Physiol.*, **75**, 57 (1970).
32 Castor, L. N., *Expl Cell Res.*, **68**, 17 (1971).
33 Hahn, G. M., Stewart, J. R., Yang, S. J., and Parker, V., *Expl Cell Res.*, **48**, 285 (1968).
34 Martin, G. S., Venuta, S., Weber, M., and Rubin, H., *Proc. natn. Acad. Sci. U.S.A.*, **68**, 2739 (1971).
35 Thrash, C. R., and Cunningham, D. D., *Nature*, **252**, 45 (1974).
36 Stoker, M., and Piggott, D., *Cell*, **3**, 207 (1974).
37 Thrower, S., and Ord, M. G., *Biochem. J.*, **144**, 361 (1974).
38 Robinson, J. H., Smith, J. A., and King, R. J. B., *Cell*, **3**, 361 (1974).
39 Sheppard, J. R., *Nature new Biol.*, **236**, 14 (1972).
40 Johnson, G. S., and Schwartz, J. P., *Expl Cell Res.*, **97**, 281 (1976).
41 Yaoi, Y., and Kanaseki, T., *Nature*, **237**, 283 (1972).
42 Igarashi, Y., and Yaoi, Y., *Nature*, **254**, 248 (1975).
43 Weiss, L., Poste, G., MacKearnin, A., and Willett, K., *J. Cell Biol.*, **64**, 135 (1975).

HOLLEY REPLIES—Growth regulation is complex, but I believe Skehan makes the subject unnecessarily confusing.

It is unlikely that all tumors are identical, so it is unrealistic to expect the absolute correlation of tumorigenicity with a single property as Skehan demands[1]. Nevertheless, there are differences between "normal" and tumorigenic cells and the two types of cells often differ greatly in their growth behaviour in cell culture. It is important to understand the differences that are observed and this was the subject of the review[2].

Skehan is influenced by what he considers to be a discovery of "growth inhibition" at low cell density[3]. In my view, Skehan has not discovered a growth inhibition but rather uses the term incorrectly, and has confused the literature. Cell cultures that begin with quiescent cells normally show a lag period, then a period in which there is one relatively synchronous cell division, and then an extended period of asynchronous growth. Skehan plots[3] such a growth curve in the form of growth rate per day and concludes that there is "growth inhibition" at the end of the period of synchronous cell division, since the rate of appearance of new cells falls. In my view, the transition from a brief period of synchronous cell division to asynchrony in a growing population is not properly called "growth inhibition."

There are several inaccuracies in Skehan's present comments[1] but they will be detected by anyone who reads the review[2].

The Salk Institute

1 Skehan, P., *Nature*, **263**, 531–532 (1976).
2 Holley, R. W., *Nature*, **258**, 487 (1975).
3 Skehan, P., *Expl Cell Res.*, **97**, 184 (1976).

reviews

Credentials of astro-archeology

A. J. Meadows

Megaliths, Myths and Men: An Introduction to Astro-Archaeology. By Peter Lancaster-Brown. Pp. 324. (Blandford: Poole, Dorset, 1976.) £4.75.

Is there such a subject as astro-archaeology? The word obviously implies not only that pre-literate cultures possessed an astronomical tradition, but also that some part of the tradition can be retrieved from surviving artefacts. The most frequently cited instance of such retrieval—and the one that forms the theme of this book—is provided by the megaliths and their alleged orientation towards astronomically significant points on the horizon. The problem is the degree of certainty that can be attached to any conclusions—after all, some of the proposed orientations, with their accompanying interpretations, are still in dispute today after nearly a century of discussion.

A central difficulty is that we are not dealing with a single debate, but with a range of topics not all of which are contested with equal vigour. At one extreme, there is the basic question —Do any astronomically significant orientations of megaliths indisputably exist? Then, somewhere in the middle, one might ask whether astronomically sophisticated orientations—for example, to particular bright stars—occur. At the other end of the spectrum, can any significance at all be attached to the so-called 'leys'—that is, alleged alignments of widely differing prehistoric features, often over considerable distances?

Mr Lancaster-Brown's book, which examines all these questions, is intended for the uninitiated. Technical astronomy is not avoided, but the necessary concepts and jargon are introduced as the need arises. A genuine beginner may find the problems of the celestial sphere occasionally hard to visualise, but the explanations are generally clear. The text is primarily concerned with present-day attitudes, but we are also given an extended survey of the historical development of interest in the subject. These two themes are by no means distinct, for part of the continuing controversy over the credentials of astro-archaeology has derived from the excessive claims of some of its past adherents. A short diversion outside

Western Europe to examine pyramidology ties in both with the main discussion of astronomically oriented monuments, and with a subsidiary question that has often been raised simultaneously: whether or not the builders of megalithic structures used a standard measure of length.

In terms of our initial range of topics, as Mr Lancaster-Brown notes, the majority of people who have examined megaliths now believe that astronomical orientation of some sort is certainly present. At the other extreme, few—except occultists—give much credence to leys. But the uncertain area remains the exact level of astronomical sophistication enshrined in the megaliths. Quite clearly, opinion in recent years has tended to allow our distant ancestors a more detailed aquaintance with astronomy than hitherto. This change has been partly

caused by the accumulation of data from new and more detailed surveys: but partly it stems from a somewhat greater willingness on the part of archaeologists to accept that our ancestors were actually capable of such sophistication. It is unfortunate in this respect that so much space in the book is devoted to Stonehenge, since, although it may be fascinating from the point of view of speculative astronomy, it is one of the astronomically over-explained monuments in the eyes of archaeologists. Again, archaeologists might wish to quibble at one or two omissions from the list of select references. But the book adequately fulfills its role as an introductory survey. □

Professor Meadows is Head of the Department of Astronomy and History of Science at the University of Leicester, UK.

Soil science and archeology

Soil Science and Archaeology. By Susan Limbrey. Pp. 384. (Academic: London.) £8.40.

IN recent years there has been a marked increase in the application of science to archaeology, but soil science has tended to be under-emphasised; this book corrects such a deficiency. The first part of the book summarises the nature of soil properties and pedogenesis. These principles are demonstrated in the following two parts by the description and evolutionary analysis of soils from selected environments. In the fourth part attention is focused on the archaeological attributes of soils.

The inevitable problem with a book of this title is defining the audience to whom it is addressed; it could have been written as a research text for soil scientists involved with archaeology or as an introductory text on soils relevant to field archaeologists. Instead it is written for archaeologists at all levels, but many will find the first part hard going, since a firm scientific background is assumed. Indeed the first three parts will prove to be of great use to students of soil science *per se;* in

particular, the chapters on the history of podzolised and chalk soils in Britain are scholarly articles in their own right.

The only general criticism which can be made is that the reader will not gain detailed advice on how to carry out various procedures—for example, the classification and mapping of soils. A common request to a soil scientist on an archaeological project is a soil or terrain map, and the fieldwork has usually to be completed within a short field season. A demonstration of the necessary techniques could have centred around the greater use of examples and illustrations.

But Dr Limbrey has produced an authoritative text which deserves a wider audience than its title might imply. The usefulness of soil science to archaeology is very well demonstrated, and the potential contribution of archaeology to our understanding of soil evolution is equally stressed.

Donald A. Davidson

Donald Davidson is a lecturer in the Department of Geography at the University of Strathclyde, UK.

Enzyme electrophoresis

Handbook of Enzyme Electrophoresis in Human Genetics. By Harry Harris and D. A. Hopkinson. (North-Holland : Amsterdam, 1976.)

RECOGNITION that enzymes may occur in multiple, electrophoretically distinct, forms first found important practical applications in clinical laboratories where attention has been concentrated on the relatively few isoenzymes known to be of diagnostic significance. The discovery of isoenzymes, however, has had an even greater impact in genetics. Enzyme heterogeneity has been studied by electrophoretic and other techniques at the MRC Human Biochemical Genetics Unit for about two decades. The authors have now provided a distillate of their vast experience in this handbook which is supplied as loose leaves in a ring binder to which supplements can be added as they become available.

After briefly outlining the significance of isoenzymes, the principles of enzyme electrophoresis and methods for their subsequent detection, techniques for the more important methods, especially starch-gel edectrophoresis, are described in sufficient detail to enable them to be set up elsewhere. The main part of the work is devoted to individual enzymes with summaries of necessary technical modifications and examples of the patterns obtained. Finally, brief tabulated sections deal with enzyme subunits and the effects of thiol reagents, together with useful appendices.

As would be expected from such distinguished authors, the work is authoritative and well written. The methods are clearly and concisely presented and, although no attempt is made to give a complete bibliography, sufficient key references are included to enable original sources to be traced. Minor criticisms concern the use of certain obsolescent terminology—for example, glutamate–oxaloacetate transaminase—and the lack of standardisation of the abbreviations. These, however, are trivial matters which do not detract from the value of the book. It will be useful, not only to geneticists, but also to clinical biochemists who at any time may encounter hitherto unsuspected enzyme anomalies.

J. H. Wilkinson

J. H. Wilkinson is Professor of Chemical Pathology at Charing Cross Hospital Medical School, University of London, UK.

Subnuclear components

Subnuclear Components: Preparation and Fractionation. Edited by G. D. Birnie. Pp. 334. (Butterworth: London and Boston, Massachusetts, March 1976.) £15.

Subnuclear Components is a valuable book for the practising research worker. It explains the immediate reasoning behind some popular preparative techniques in an area in which the end justified the means. Happily, assessment of the end-product is included in most of the chapters. The discussions all lie in the areas between detailed descriptions of techniques and expositions of general principles, and provide a very valuable insight into the thinking and background associated with some important experimental procedures. Each of the eight chapters is written by a different author or group of authors, all of whom are busy research workers themselves; and it is not surprising therefore that the text and proof reading are not as polished as I would

otherwise expect for £15 per copy. Less excusable is the long gap between writing and publication, during which time some material has been left behind by events such as the nucleosome model for chromatin; and quite crazy is the tiny print used for the figure captions. Although many of the main techniques have been covered there are gaps in such areas as ribonucleoprotein particles, mitotic apparatus and DNA fractionation (DNA preparation is well covered in a chapter by Dr Butterworth). In the best chapters the available methods are critically discussed and a few chosen for detailed analysis. There are, however, some important gaps—for example, in the otherwise excellent discussion of histone preparation, chromatographic methods are almost ignored. Such problems are inevitable when the authors are presenting the detailed 'behind-the-scenes' analysis which make *Subnuclear Components* so valuable. **Harry R. Matthews**

Harry R. Matthews is a senior lecturer in the Department of Physics at the Portsmouth Polytechnic, UK

Protoplasmic connections in plants

Intercellular Communication in Plants : Studies on Plasmodesmata. Edited by B. E. S. Gunning and A. W. Robards. Pp. xvi + 387. (Springer : Berlin and New York, 1976.) DM72; $29.60.

THIS volume contains 14 review papers presented and discussed at a meeting held in Canberra in June 1975. The aim of the volume according to the editors is to summarise what is known about the nature of plasmodesmata and assess the significance of these intercellular protoplasmic connections in plant physiology. The contributors have performed a valuable service in collecting together information which was previously scattered widely in the literature, and the editors are to be congratulated on the skill with which they have integrated the papers and summarised the discussions at what must have been a very lively meeting.

It is now nearly 100 years since Tangl described strands of protoplasm traversing the cell walls in the endosperm of certain seeds, but because of their extreme fineness (30–60 nm) their very existence remained in some doubt until the advent of the electron microscope. As Robards points out in this volume, in spite of intensive study in the past 10 years, many of the details of struc-

ture remain unresolved. In particular the relationship between the desmotubule and the endoplasmic reticulum of the adjoining cells is still uncertain.

Even greater controversy surrounds the operation of plasmodesmata and much of the volume is devoted to this problem. The early suggestions of Haberlandt that protoplasmic connections may function both in the translocation of substances and transmission of stimuli are supported by a wealth of more recent circumstantial evidence, but an unequivocal demonstration that things actually move through plasmodesmata is still awaited. Tyree has established that plasmodesmata may have sufficient capacity to transport solutes at the fluxes observed, but the basis of such transport is still obscure. This volume, with its extensive bibliography, will form an excellent springboard for anyone wishing to advance our knowledge in an intriguing field of research.

The book is printed from 'camera-ready' typescripts without justified margins and one might have expected that the saving in time and cost would be reflected in more rapid publication and lower price than has been achieved. **James F. Sutcliffe**

James Sutcliffe is Professor of Plant Physiology in the School of Biological Sciences at the University of Sussex, UK, and Editor-in-Chief of Annals of Botany.

obituary

John Richardson Marrack, DSO, MC, Emeritus Professor of Chemical Pathology in the University of London, died in the USA on June 13, 1976. He was born on November 26, 1886 at Clevedon, Somerset, but soon moved to Tiverton, where he attended Blundell's School, to which he remained greatly attached, and of whose Old Boy's association he later became Vice-President. He went to St John's College, Cambridge and then to the London Hospital Medical College, graduating in 1908.

His first research was on rheumatoid arthritis as a John Lucas Walker student, and later Beit Memorial Fellow at the laboratories of the Cambridge Research Hospital (which has become the Strangeways Laboratory). After the First World War, when he served in the RAMC, he went to the London Hospital as lecturer in Chemical Pathology. He became interested in the properties of colloids, initially from studying the binding of calcium by serum proteins, and came to the conclusion that colloid interactions were caused by definable and verifiable physical and chemical forces, acting between distinct protein entities. As an example he chose antibodies, whose nature was quite unknown and whose very existence as separate entities was doubted. In 1930 he showed that diphtheria antitoxin behaved as a distinct protein whose interaction with diphtheria toxin could be measured quantitatively. In a monograph published in 1934 (*The Chemistry of Antigens and Antibodies*) he proposed that the specific affinity of antibodies for antigens is determined by the same factors which determine the specific binding of molecules to form crystals, that is, the shape of the molecules and the spatial distribution and strength of

polar forces. The monograph contains a clear diagram elaborating the theoretical studies of Heidelberger and Kendall to illustrate what has now become accepted as the 'lattice hypothesis' of antigen–antibody interactions.

Revised in 1938 this work has had a lasting influence, and convinced many chemists and biochemists that immunology was a fit subject for scientific study by themselves as well as by bacteriologists and serologists. Marrack was also the first to use methods which are now commonplace: equilibrium dialysis, whereby he indicated that anti-hapten antibodies were probably bivalent, and the attachment of coloured dyes to antibacterial antibodies, which inspired Albert Coons later to develop the technique of immunofluorescence. Marrack wrote few papers, by present day standards, and his encyclopaedic knowledge of immunochemistry appeared mostly in review articles.

John Marrack was a colourful character. Behind a shyness and apparent abruptness lay kindness and intellectual integrity. He always wanted to be an athlete and was by temperament a fighter—for seven years he was welterweight champion in the London University boxing tournaments. Throughout the whole of his adult life he was a keen walker (he knew Dartmoor intimately) and he never drove where he could go by bicycle. On more than one occasion when roused to righteous anger he took the law into his own hands and used his fists: once to apprehend a thief in the laboratory and again to despatch a gang of hoodlums who misguidedly attacked him on Whitechapel Station. His war record in the RAMC—DSO as a line medical officer and MC for investigations on the poison gas used

against the British Army in 1917—illustrates this aspect of his character. So also does his consistent championship of the underdog.

During the Civil War in Spain, he was an active member of the Spanish Medical Aid Committee, and visited the International Brigade and the Spanish Republican army. About this time he became deeply concerned about the nutrition of children in Britain, influenced by L. J. Harris and Jack Drummond, and spent much time and effort campaigning for the Children's Nutrition Council—to such good effect that he was adviser to the Ministry of Food during the Second World War and wrote in 1942 a book (*Food and Planning*) which influenced the post-war planning of nutrition.

These activities were regarded by many of his contemporaries as indicating that he was finished with research, but were entirely consistent with his character. When he returned to the laboratory in the Department of Pathology at Cambridge in 1952, his main work had, in fact, been completed, but he began to exploit the growing knowledge of the structure of antibodies while devoting most of his energy to editing, for its first ten years, virtually single handed, the new journal *Immunology*. By now the importance of his earlier work had become widely understood and recognised and at the age of 76 he was made Visiting Professor at the University of Texas. At the First International Congress of Immunology in 1971 he was one of five to receive the Distinguished Service Award "For revolutionary ideas that have become commonplace in his lifetime, and for pioneering work in the physicochemical interpretation of antigen–antibody interactions".

J. H. Humphrey

Marie Laura Violet Gayler, March 1891–August 1976, was the youngest of five daughters of Mr William Gayler, Director of Stamps and Excise at Somerset House. Her mother was an artist, a Gold Medallist of the Slade School, whose paintings were often exhibited at the Royal Academy. Marie graduated from London University in 1912, and after teaching botany at the Colstan Girls' School, Bristol, in 1915 she joined Walter Rosenhain's scientific staff in the Metallurgy Department of the National Physical Laboratory, and later married Dr J. L.

Haughton, a member of the same Department. She retired in 1947.

Marie Gayler and a physical chemist, Miss I. H. Hadfield, were the first women to be appointed to the scientific staff of the Department. She became a distinguished member of Rosenhain's team, which in the 1920s and 30s at the NPL, helped to lay the scientific foundations of physical metallurgy and to give this country the leading position which it enjoyed in the subject for a couple of decades or more.

Marie Gayler's outstanding contribution, with Hanson and Haughton,

was the elucidation of the mechanism of age-hardening in the duralumin family of aluminium alloys, which had been developed empirically by Wilm in Germany. A very important outcome of the NPL work was Y-alloy, an aluminium alloy which contained nickel as well as copper, magnesium and silicon, the normal alloying elements in duralumin. The presence of nickel greatly improves the strength and hardness of age-hardened duralumin at temperatures of 150–200 °C. This makes Y-alloy eminently suitable as a material for the pistons of internal

combustion engines. Many derivatives of Y-alloy have been developed, among which are the RR series of alloys, one of which is used in the skin of the Concorde which is heated by passage through the air at supersonic speeds.

In the 1930s and 40s Marie Gayler took over the work at the NPL on dental amalgams. To understand the mechanism of the setting and hardening of amalgams she followed the traditional NPL procedure, which was to establish the metallurgical constitution of the alloys, consisting essentially of silver, tin and mercury. The setting proceeds by diffusion at room temperature of the mercury into the powdered silver-tin alloy. Novel metallographic techniques had to be worked out for the study of the complex processes of diffusion and reaction. To ensure the satisfactory prosthetic performance of the amalgam filling, close control of the volume changes occurring during and after setting must be assured. In recognition of this side of her work Marie Gayler was made an Honorary Member of the British Dental Association in 1947.

Although her principal researches were concerned with aluminium alloys and dental amalgams, Marie Gayler also carried out important investigations on iron-manganese alloys, and on the melting points of pure silicon and iron. In 1947 the Institute of Metals, in whose Journal many of Marie Gayler's papers were published, awarded its Platinum Medal jointly to her and her husband, Dr J. L. Haughton.

After her retirement she was able to devote more time to her interest in sculpture; she sculpted the head of the late Professor Hume-Rothery; which now stands in the library of the Department of Metallurgy in Oxford University.

Marie Gayler readily absorbed and whole-heartedly transmitted Rosenhain's enthusiasm and scientific resourcefulness. She worthily maintained and enhanced the traditions of the Metallurgy Department of the NPL, and her surviving colleagues remember her as a warm and helpful colleague of striking and attractive appearance.

A. J. Murphy

announcements

Meetings

November 11, **Ecdysone Workshop,** London (Dr. D. L. Whitehead, Tsetse Research Laboratory, Department of Veterinary Medicine, Langford House, Langford, Bristol BS18 7DU, UK).

January 29–30, 1977, **Human Skin Banking,** Milwaukee, Wisconsin (Ralph M. Guttman, MS, Director, St Mary's Skin Bank, PO Box 503, Milwaukee, Wisconsin 53201).

March 2–4, 1977, **Cell Differentiation and Neoplasia,** Houston (Stephen C. Stuyck, MD, Anderson Hospital and Tumor Institute, Houston, Texas 77030).

March 28–April 1, 1977, **Scanning Electron Microscopy,** Chicago (Deadline for abstracts: October 25) (Om Johari, Annual SEM Symposia, IIT Research Institute, 10 W 35th Street, Chicago, Illinois 60616).

May 3, 1977, **Crop Protection,** Ghent (Prof. Ir. R. H. Kips, Chairman of the Organising Committee, International Symposium on Crop Protection, Faculteit van de Landbouwwetenschappen, Coupure Links 533, B-9000 Ghent, Belgium).

May 17–21, 1977, **Study of Macromolecules by NMR,** Grasmere, UK (Dr D. H. Richards, Explosives Research and Development Est., Non-metallic Materials Branch, Powdermill Lane, Waltham Abbey, Essex).

July 11–14, 1977, **Gas Kinetics,** Manchester (Fifth International Symposium on Gas Kinetics, UMIST, PO Box 88, Manchester M60 1QD).

July 11–15, 1977, **Organometallic and Co-ordination Compounds of Ger-**

Person to Person

A limited number of copies of the *Genetic Hazards to Man from Environmental Agents,* the proceedings of a symposium held in Ottawa in May, 1975, and published in *Mutat. Res.,* **33,** (1), (1975) are available to scientists who would not have access to it otherwise. Please write to Dr John A. Heedle, Department of Biology, York University, 4700 Keele Street, Downsview, Ontario M3J 1P3, Canada.

I am conducting experiments on the variations of fungal pathogens of rice induced by pesticides. Those who are working on similar topics may please contact me if interested in exchanging views on this subject (Dr S. Balakrishnan, Dept of Pathology, College of Agriculture, Vellayani PO, Trivandrum, Kerala, India).

There will be no charge for this service. Send items (not more than 60 words) to Martin Goldman at the London office. The section will include exchanges of accommodation, personal announcements and scientific queries. We reserve the right to decline material submitted. No commercial transactions.

manium, Tin and Lead, Nottingham (Dr P. G. Harrison, Secretary, 2nd Ge, Sn and Pb Conference, Department of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, UK).

August 8–19, 1977, **Physics of Quantum Electronics,** Telluride, Colorado (Professor S. F. Jacobs, Rt. 2, Box 732D, Tucson, Arizona 85715).

August 21–26, 1977, **Singlet Oxygen and Related Species in Chemistry and Biology,** Pinawa (D. Fundytus, Conference Secretary, Technical Information Services, Whiteshell Nuclear Research Establishment, Atomic Energy of Canada Ltd, Pinawa, Manitoba, Canada, ROE 1LO).

August 30–September 3, 1977, **Fourth European Crystallographic Meeting,** Oxford (Dr C. K. Prout, Chemical Crystallography Laboratory, 9 Parks Road, Oxford OX1 3PD, UK).

September 5–10, 1977, **Biology of Connective Tissue,** Uppsala (Professor T. Laurent, Biomedical Centre, University of Uppsala, Box 575, S-751 23 Uppsala, Sweden).

September 12–16, 1977, **Bioindicatores Deteriorisationis Regionis,** Liblice, nr Prague (Ústav krajinné ekologie CSAV honice, Czechoslovakia).

September 25–October 11, 1977, **Kimberlite,** Santa Fé, New Mexico (2nd (Ing. J. Spálény CSc.), 252 43 Prů-International Kimberlite Conference, Sylvia-K Inc., 5671 Blue Saga Drive, Littletown, Colorado 80123).

October 11–13, 1977, **Oceanic Fronts,** New Orleans, Louisiana (Deadline for abstracts: November 1) (American Geophysical Union, 1909 K Street, N.W., Washington, D.C. 20006).

November 14–18, 1977, **Pan American Conference on Forensic Applications of Medicine, Dentistry, Pathology and Palaeopathology,** Mexico City (Dr William G. Eckert, Laboratory, St Francis Hospital, Wichita, Kansas 67214).

More than facts, judgments

THE idea of a Science Court has gained a modest amount of momentum in the United States (*Science*, August 20, p 653; *Nature*, October 7, p 454). It would concern itself solely with providing "factual statements of the highest presumptive validity consistent with time constraints". For this, an adversary procedure would be used and a panel of judges with skills in adjacent areas would be expected to examine statements and, if necessary, write their own opinions on contested matters. Issues which might come early to such a court are: should fluorocarbons be banned because of their effect on the ozone layer? Should water supplies be fluoridated? Should a specific nuclear plant be licensed?

The working party's interim report (in *Science*) has the clear implication in its opening paragraph that a Science Court, by sticking closely to what can be factually established, will be able to hand over something useful to the decision-makers of society, who will then incorporate social value questions. A scientific fact is defined as a result, or an anticipated result, of an experiment or an observation of nature. On the other hand, the report continues, "we have no illusions that this procedure will arrive at the truth, which is elusive and tends to change from year to year." Thereby the science court neatly ducks the issues and relegates science to a backroom, purely technical rôle.

The scientist is most unlikely to be able to deliver to the decision-maker any useful sort of factual statement, because he is hardly going to be allowed to perform the appropriately large experiment or observation. All he can generally supply in the way of facts is some results from pilot projects, some calculations which may be relevant and so on. What the good scientist should also be competent to provide, however, is inference, and this, albeit tentative and hedged-about, is what the decision-maker needs and what the science court seems to avoid.

Factual statements of the highest presumptive validity would merely be about rats, about rocket samples, about tensile strengths. Those involved in public policy need to know whether, in the scientist's best judgment, such statements can be generalised. Intelligent customers for these sorts of judgments know full well that scientific 'truth', being a whole level higher than facts, is often every bit as elusive and changeable as political and economic 'truth'. But they still expect the scientist to go beyond the solid ground of his facts.

It is difficult to see that a science court has anything new to offer if it steers clear of this territory—indeed the exercise could be positively dangerous in that others, less qualified, may be tempted to draw their own conclusions from such handily accumulated and uninterpreted facts. □

More than judgments, action?

THE best debates never die. As the arguments about nuclear power continued in Britain last week with the resumption of the SGHWR enquiry by the House of Commons Select Committee on Science and Technology, life was again being breathed into a still more fundamental (though still less newsworthy) debate—the centuries-old argument about the depletion of the Earth's natural resources. It came with the publication of *Materials and Energy Resources*, the report from a three-year-old working party on resources embraced by the UK Institution of Chemical Engineers (I.Chem.E.).

At £5 it is a costly two-score pages, not helped by the I.Chem.E. wanting, apparently, to have its cake and eat it. It is both seeking to stimulate reasoned discussion on crucial resource problems and trying to get something done about them. Thus, while the report rightly points out the vital need to see materials and energy problems as two sides of the same coin, and traces the broadly-based working party's own valuable assessments of these problems, it is also being presented as an attempt to influence policy. Not, though, general resources policy internationally, but specifically, and as a matter of strategic importance, UK government policy.

That is laudable enough, perhaps, but represents a curious disposition for such a study from such a body. After all, the disciplines it represents and the problems it is tackling are nothing if not international, as are the phenomena like international cartels which it acknowledges. Since it also focuses its attention on a period 10 to 35 years hence (by which time, if the past is anything to go by, the world order ought to have changed), the working party might have done better to look more closely for international solutions.

Not that it has done wrong in sending the report to the Secretaries of State for Energy, Industry and the Environment. But, in spite of what was described enthusiastically last week as a positive response from Industry, the fears being expressed privately that the report could easily lose itself in the bureaucracy are more than justified. The I.Chem.E., which probably wants simply to be a professional group turning out professional work, ought to leave both this problem and the job of actually campaigning for changes in policy to the pressure groups.

The present importance of the nuclear issue, in the UK and elsewhere, is not a little due to fringe groups making effective use of the information and judgment from reputable bodies ill-placed if not disinclined to deploy it themselves. The resources issue ought to be next. The best debates never die. □

Seveso: the aftermath

Alastair Hay reports on developments since a
chemical explosion rocked this northern Italian town

Four-year-old Séveso victim

THREE months have elapsed since the now notorious Séveso accident of July 10. It occurred when a reactor producing trichlorophenol at the ICMESA chemical plant—owned by Givaudan, a Hoffman-La Roche subsidiary—became overheated: the rapid pressure increase that followed led to the discharge of its contents through a safety valve directly into the atmosphere. Scientists have since been engaged in decontamination work in the area, which was affected by the fallout of TCDD (2,3,7,8-tetrachlorodibenzo-para-dioxin) as the disastrous consequence of that discharge. A considerable amount of disturbing information has now come to light concerning both the operation of the ICMESA plant, and the decontamination procedures adopted.

Givaudan began to produce trichlorophenol at the ICMESA plant only in 1975. The monthly output of the plant was 5 tons but, in producing this quantity of trichlorophenol, the reaction also produced 1.2 kg of dioxin. This, according to an Italian trade union document, is the quantity which ICMESA was required to dispose of every month. It represents a high level of dioxin formation, of the order of 200 p.p.m.

According to a report in *The Times* earlier this week, however, Dr Donald Lee, the British expert now investigating the accident, is thought to have told the Italian authorities that up to 130 kg of dioxin could in fact have been produced in the explosion. His recommendations are said to include the suggestions that residents should have medical tests for the rest of their lives, that the contaminated area should be monitored for decades, that this and a buffer area should be sealed off to all but authorised personnel, and that decontamination procedures should involve the creation of forests and dismantling of buildings rather than the shifting of earth and the washing of buildings.

Although a senior Roche executive says, according to the Italian weekly *L'Espresso*, that the company was producing TCDD deliberately for military purposes, Givaudan has denied the allegation. But Professor Fritz Mori, designer of the trichlorophenol reactor at ICMESA, has referred to the cause of the dioxin formation through a sudden rise in temperature within the reactor and gone on record as saying it would be impossible "unless the reactor was being used to produce substances other than the trichlorophenol it was designed to produce". The design of the reactor involved in the accident has come in for severe criticism, for incorporating safety features inadequate for the production of chlorinated phenols.

Information possibly vital to the resolution of the problems of plant operation, design and safety, however, is still not available to the Italian authorities. As for the estimate that the residents of Séveso and its surrounding area will not be allowed to return to their homes for at least a year, this is regarded as optimistic by many scientists. Although vegetation from the less contaminated zone is now being incinerated, no specific plan for removing dioxin from the soil has yet been agreed. Originally, topsoil from the total area contaminated was to have been cleared and processed to remove dioxin. However, some recent tests have revealed dioxin at a depth of 25 cm. The original proposition, which relied on TCDD being close to the surface, is now considered to be too difficult an undertaking, and operations will be confined to removing dioxin from the 30–40 hectare zone most severely contaminated.

Antidote claimed

Givaudan claims to have developed an antidote which will destroy the dioxin, consisting of an olive oil–water emulsion and cyclohexanone. In principle, the dioxin is detoxified by rupture of the linked ring system and the removal of the chlorine atoms by the unsaturated bonds of polyunsaturated fatty acids present in the olive oil. Ultraviolet light absorbed by the cyclohexanone initiates the reaction sequence. However, the Italian authorities, as well as American and British scientists, are sceptical about the success of the Givaudan claim. The Italians and Americans maintain that there is no effective antidote for TCDD. And Professor Derek Bryce-Smith of Reading University says sunlight does not contain a high proportion of ultraviolet in the wavelength range where cyclohexanone absorbs it. Although the photochemical reaction may take place in the laboratory, there is, he says, a real danger of moving dioxin so that it is harder to get to.

TCDD is virtually insoluble in water, and for this reason is not translocated rapidly in soil. Placing a dioxin solvent on the soil—TCDD is partially soluble in organic solvents—could obviously assist in moving the dioxin to greater depths. Another complication is the fact that cyclohexanone is itself a dangerous chemical: its flash point, only 63 °C, is well below its boiling point of 155.6 °C; in terms of human toxicity, it is considered a weak narcotic, and it also irritates skin and mucous membranes; the lethal dose for mice is 8,000 p.p.m. in air.

The reports causing most worry, though, are those concerning the results of the blood tests carried out on the ten thousand residents of the Séveso area. Information not yet released by the Italian health officials confirms that 5–10% of the people examined are showing reduced white cell blood counts. Although it is only three months since dioxin was released, and a reduction in blood lymphocyte count is a secondary effect of damage to the immune response tissue—which dioxin is known to affect—these results are causing great concern. The situation is expected to worsen, and more residents may show adverse clinical symptoms; persons already affected could record a further reduction in their white cell count. But some scientists who have recently returned to the UK after visiting Italy feel that symptoms of dioxin poisoning are less serious than was originally feared.

Given the known teratogenic properties of dioxin, 150 women from the Séveso area who were in the first trimester of pregnancy at the time of the explosion have applied for an abortion. By the end of September, 25 women had received government authorisation for the operation, in the face of great opposition from the Catholic Church. Authorisation for another 125 women is also likely.

Health officials have examined a total of 730 pregnant women formerly resident in the contaminated area. The difficulties in securing legal abortions have added considerable impetus to the campaign for abortion law reform; the Italian parliament is to consider several bills advocating changes.

In view of the known toxicity of TCDD and the knowledge that it is a constant by-product of the trichlorophenol manufacturing process, many scientists are now beginning to question the need for the continuation of trichlorophenol production. One of the two

The possible alternatives

THE available alternatives to 2,4,5-T are 'Amcide', 'Glyphosate' and 'Krenite'; all three are effective brushwood killers. Amcide has been available in the UK since the early 1960s; Glyphosate is a more recent addition, and Krenite is at present only on sale in the US and West Germany.

These alternatives, when compared with 2,4,5-T, have some definite advantages. Amcide, made by Nissan in Japan, has an LD_{50} value for rats of 3,900 mg kg^{-1}; Glyphosate, made by Monsanto, has an LD_{50} value of 4,900 mg kg^{-1}. These toxicity ratings are well below that for 2,4,5-T (LD_{50} value 300 mg kg^{-1}). Krenite, made by Dupont, has an LD_{50} value of 24,000 mg kg^{-1}, a toxicity rating one eightieth of that for 2,4,5-T.

The reaction to produce Amcide presents none of the potential hazards associated with trichlorophenol manufacture in terms of toxic by-products. Information has not been made available for Glyphosate or Krenite production. However, in the case of Krenite the US Environmental Protection Agency (EPA) considers its manufacture to conform to accepted safety standards, and in view of its comparatively low toxicity has declared it as safe for use even on land adjacent to domestic water supply reservoirs and streams.

The disadvantages associated with the three alternatives are primarily those of cost. Amcide and Glyphosate are more expensive than 2,4,5-T when the concentrations of herbicide necessary to effect the same plant kill ratio are considered; Glyphosate is nearly five times as costly. Some potential buyers of Krenite consider that it, too, may be expensive when it is introduced on the UK market. The only other problem of any consequence relates to Amcide. In contrast to 2,4,5-T which is absorbed through leaves following spraying, Amcide must be applied to cut surfaces on plants. Its method of application is therefore considerably more labour intensive.

Hexachlorophene, the other major product derived from trichlorophenol, is a general poison effective in the control of gram-positive bacteria. In the cosmetics industry hexachlorophene is used as a preservative. For medical purposes hexachlorophene is used in the control of staphylococcal organisms. The bactericide has four main uses: treatment of acne and impetigo, cleansing of intact skin around burns or wounds, pre-surgical washing, and cleansing of new-born infants, particularly the umbilical cord.

Its use, however, has been much reduced, and the industry is reported as being able to dispense with it altogether. Chlorhexidine—a bactericide now cleared for sale in the US—is used surgically in the UK as a skin cleanser, wound steriliser and for pre-surgical washing. On the question of the cleansing of the newly-born infant, however, there is a difference of opinion as to which of the two bactericides is the most effective.

Maternity clinics and nurseries are particularly open to bacterial cross-infection by the staphylococccal organisms. One of the most common is *S. aureus*. In the 1940s this organism was responsible for frequent epidemics in nurseries, with a consequent increase in infant mortality. When hexachlorophene was first marketed by Givaudan in the late 1940s it proved to be effective both in the routine containment of *S. aureus* and in controlling the organism in the case of an epidemic. As a result of its efficiency, hexachlorophene rapidly replaced the bactericide 'Triple Dye' in use at that time, to become the most widely used anti-bacterial agent in nurseries.

Two events in 1971 and 1972, however, caused users to reconsider their judgment. The first was a report by the EPA showing hexachlorophene to cause oedema and hindlimb paralysis in experiments on mice. The second was the death in France of 35 infants following the use of talcum powder containing 6% hexachlorophene. The neurological damage which led to the death of the children was caused by a twenty-fold increase of hexachlorophene concentration in the talc, the result of a manufacturing error.

Many maternity units in the UK have since reduced the amount of hexachlorophene used for infant washes. Currently less than half use hexachlorophene at all. Others rely on alcohol, used either alone or with chlorhexidine. Consultants at one cross-infection laboratory now recommend nurseries to avoid the use of hexachlorophene altogether for routine washes. The reasoning behind this recommendation lies in the evolution of *S. aureus* itself, which has evolved through several forms since the 1940s and is now active as a complex of *S. aureus*.

The present generation of the bacterium is not causing serious epidemics in British hospitals which are fatal to children. Indeed the risk associated with hexachlorophene use is considered to present a greater threat than that represented by the staphyloccus itself. A further consideration is the fact that many British maternity units use a concentration of hexachlorophene to control *S. aureus* which is too low to kill the bacterium.

In the routine control of *S. aureus* chlorhexidine is as effective as hexachlorophene, and has the added advantage of an LD_{50} value ten times higher. Chlorhexidine is also reported to be just as efficient in controlling local epidemics of the present milder strain of the bacterium. It has not been tested, however, with more lethal strains of the staphylococcus, whereas hexachlorophene has been shown to control dangerous epidemics of *S. aureus* in the past. Some bacteriologists anticipate that chlorhexidine will prove to be equally effective, but are reluctant to recommend it unequivocally as an alternative until this has been demonstrated conclusively.

Although the principal purpose behind trichlorophenol production is the synthesis of 2,4,5-T and hexachlorophene, the chlorinated phenol is still used to some extent as a slime control agent in the paper-making industry. It has never been as popular as pentachlorophenol, used for the same purpose, and what use it has had has been much reduced, due mainly to the criticisms of the unrestricted industrial use of polychlorinated biphenyls (PCBs) such as DDT, advanced by the environmental lobby.

In common with many industries which discharge effluent directly into rivers, the paper manufacturers have been particularly sensitive to the charge that their plant operating procedures are damaging to the environment. Because of the known risks associated with extensive PCB use, paper manufacturers in the UK have considerably reduced their use of all chlorinated phenol derivatives. Agents now preferred for the prevention of fungal growth include methylene *bis* thiocyanate together with some organobromine and organosulphur products.

products derived from trichlorophenol —the herbicide 2,4,5-trichlorophenoxy-acetic, 2,4,5-T (the other is the bactericide hexachlorophene) — has many formulations which contain the dioxin contaminant. In view of this the Norwegian authorities banned the use of the herbicide altogether in 1973. The US Environmental Protection Agency (EPA) regards 2,4,5-T as a product which may be too hazardous to man and the environment to permit continued use.

In the UK, however, the Ministry of Agriculture, Fisheries and Food's 1976 "List of Approved Products for Farmers and Growers" contains five formulations of 2,4,5-T alone, and eight of 2,4,5-T in combination with 2,4-D (2,4-dichlorophenoxyacetic acid). These products are sold as bramble, brushwood and nettle killers. A new formulation of 2,4,5-T, 'Silvapron T', produced by British Petroleum, is currently on trial in Britain. The herbicide is dispersed in an oil-based mixture which

The cost mounts

avoids some of the problems of volatilisation encountered with the traditional water-based mixtures. This BP formulation is restricted to forest areas.

Neither product unique

But neither 2,4,5-T nor hexachlorophene is unique. Alternative herbicides and an equally effective bactericide are available (see box). Because of this, many scientists now question the desirability of using a reaction as dangerous as that involved in synthesising trichlorophenol. They point out that, with one exception, every plant producing trichlorophenol has had a serious accident involving release of dioxin, and injury to workers and the general public. The list of accidents begins with Monsanto (US) in 1949; then comes Badische Anilin und Soda Fabrik AG (West Germany) in 1953, Philips Duphar (Holland) in 1963, Dow Chemical Company (US) in the early 1960s, Coalite and Chemical Products Ltd (UK) in 1968, and finally, Givaudan-ICMESA (Italy) in 1976. Bayer of Leberkusen (West Germany) is the sole apparent exception.

But it is not only scientists who are concerned about the operation of the trichlorophenol process. The Italian authorities have invited Givaudan to build a new chemical plant near Séveso, with the provisos that it utilises reaction processes known to be safe and that trichlorophenol is not produced. Two other manufacturers of trichlorophenol are equally concerned; the Coalite and Bayer plants have both temporarily ceased operation. Additional safety features are to be introduced at Coalite as a result of recommendations by the UK Health and Safety Executive, but forty-nine of the fifty workers have refused to work there under any circumstances.

The trichlorophenol manufacturing process is only one of the potentially

hazardous operations which will now come under closer scrutiny as a result of the devastation at Séveso. There are few processes which can have the unique record of an almost one-to-one ratio of accidents to production sites. However, there may be some where the value to society of the products manufactured is not sufficient to justify continued operation. The disaster in Italy will provide the backcloth for the debate which could help to resolve these issues.

● Yet another release of a dangerous chemical occurred in Italy on 26 September. This time between 10 and 30 tonnes (according to different sources) of arsenious oxide "escaped" from the ANIC chemical plant—a subsidiary of the Italian state fuel concern ENI—near Manfredonia on the Adriatic coast.

Over fifty people including two infants have been hospitalised with suspected arsenic poisoning (LD_{50} for rats is 138 mg kg^{-1}). The ANIC plant management only admitted the risk to local residents when the results of tests carried out by health officials confirmed a serious pollution hazard. The mayor of a town in the affected area appealed to the Italian government for assistance; no response came for several days. ANIC officials now say all but 1 km^2 of the area has been rendered harmless.

This has been done by spraying with calcium chloride and ferric sulphate solutions. When arsenious oxide is in a soluble form these two additives will form complexes of calcium pyro-arsenite and ferric enneaoxyarsenite respectively. Both complexes are considerably less soluble than the oxide and would thus reduce the amount of arsenic leaching into the water system and spreading far beyond the present area of pollution.

Feeding cows to be destroyed

Inspecting ANIC plant

USA

Nuclear power's litmus test

A crucial test of public confidence in nuclear power will take place in six states during next month's Presidential election. Colin Norman reports from Washington.

WHEN residents of Arizona, Colorado, Montana, Ohio, Oregon and Washington mark their ballot papers on November 2, they will do more than cast votes for candidates for sundry political offices. They will also vote on resolutions which would place strict—some say crippling—controls on nuclear power plants in their states.

The resolutions are stirring up considerable debate, and large sums of money are being spent both to promote and defeat them. Although the number of power plants affected is relatively small compared with the total number planned for the United States, the votes are regarded as a litmus test of public attitudes towards nuclear power. With the election less than a month away, opinion polls indicate that the resolutions stand a good chance of being approved in at least two states—Colorado and Oregon.

Although there are some differences between the resolutions on the ballots in the six states, they would all set three important conditions on the operation of nuclear power plants.

● There must be no limit to the total amount of damages which could be claimed by victims of a nuclear accident. At present, a federal law, the Price-Anderson Act, limits the nuclear industry's liability to a total of $560 million for each accident.

● Before new power plants could be operated, the state legislature would have to be convinced that major safety systems—including the much-discussed emergency core cooling system—would operate properly in an emergency. Approval would require a two-thirds vote in the legislature.

● Similarly, at least two thirds of the legislature would have to be satisfied that adequate, proven technologies exist for transport, storage, and disposal of radioactive wastes.

At present, it is doubtful that the nuclear industry could meet those conditions. Thus, passage of the resolution in any state would virtually rule out nuclear expansion in that state.

Voters in California were presented with a similar choice last June, when a nuclear safety proposition was placed on the ballot papers in the Presidential primary elections. That proposition, which would have imposed the same three conditions on nuclear plants in

California, was eventually defeated.

Supporters of the resolution in Oregon and Colorado, and to some extent in the other four states, expect to fare better, however. For one thing, the steam was taken out of the California campaign at the last moment when the state legislature passed a package of nuclear safety bills which included some of the proposition's requirements. But, more important, the California proposition would have applied to existing, as well as planned, power plants. The Oregon and Colorado measures, by contrast, would specifically exempt existing plants.

A potentially important factor in the vote on the Oregon resolution is that during the Oregon primary election earlier this year, Jimmy Carter, the Democratic Presidential candidate, virtually endorsed the resolution. He said that although he would not try to advise people how to vote on the matter, if he were an Oregonian, he would support the resolution. He did not endorse the California proposition, however, largely because it would have

applied to existing plants.

Supporters of the resolution also like to point out that both Colorado and Oregon have approved a number of environmental laws long before other states. Oregon, for example, was the first state to outlaw the sale of plastic bottles and aerosol sprays containing fluorocarbons, and among Colorado's environmental achievements is the passage of a resolution which stopped the Winter Olympics being held there.

But the most solid piece of evidence that the nuclear industry is in trouble in Colorado and Oregon comes from public opinion polls published early in October in newspapers in Denver and Portland. They suggested that public sympathy was then running nearly two to one in favour of the resolutions. Although that margin is sure to close as the election draws near industry officials say they are very concerned.

Thus, if the resolutions are adopted by any state this November, it would give the anti-nuclear movement a boost in future efforts. But, by the same token, if the resolutions are defeated, the nuclear industry could claim a vote of confidence. Clearly, the stakes are high.   □

GAO knocks ERDA

Although the Energy Research and Development Administration (ERDA) is responsible for developing nuclear energy technologies in the United States, ERDA officials insist that they have not tried to influence the outcome of the referenda on nuclear safety being held in various states. Those assertions are, however, called into question by a report published last week by the General Accounting Office (GAO), an investigative agency of the US Congress.

According to the report, between February and April 1976—when debate on the California referendum was raging—ERDA distributed 78,600 copies of a pro-nuclear pamphlet in California. According to GAO, the pamphlet "was not objective, is propaganda, and was not a proper document for release to the public". The pamphlet was used by the nuclear industry in California in their campaign to defeat the California proposition.

ERDA officials have claimed that the document—a series of so-called myths and facts about nuclear power—was prepared solely for internal distribution within the agency. It was intended, they have argued, for employees in the controversial liquid metal fast breeder reactor pro-

gramme, to help answer criticisms of the programme and to lift morale.

GAO found such assertions hard to swallow, however, since the fast breeder programme employs only about 6,700 people, yet 100,000 copies of the pamphlet were printed. Moreover, GAO found it difficult to understand why the bulk of the copies were sent to California, while most of the breeder reactor work is going on in other states.

Nevertheless, GAO stopped short of actually accusing ERDA officials of trying to influence the California vote—in fact, it specifically states that it could find no such evidence and suggested that ERDA officials had not actually violated any laws. The report did not attempt to explain why so many copies of the pamphlet were distributed in California at that particular moment, however.

As for the quality of the pamphlet, the GAO report suggested that it was so superficial and misleading that it was not even suitable for distribution to ERDA employees as part of morale lifting effort. "ERDA should not place itself in the position of misleading others—whether it be the public or its own or contractor employees—for the sake of raising morale", the report argued.

CANADA

A matter of judgment

David Spurgeon writes from Ottawa on a report covering Canada's problematic science-industry relationship

IN 1973, the Ministry of State for Science and Technology (MOSST) introduced a "Make-or-Buy" policy, the aim of which was to increase the proportion of government-funded research and development (R&D) carried out under contract by industry as opposed to government laboratories. The government hoped to strengthen the innovative capacity of Canadian industry and to enhance its competitive position. The policy has not generally been regarded as a howling success, especially by industrialists. In the year following its introduction, in fact, the Cabinet found it necessary to expand the original policy to allow financing of unsolicited proposals for R&D from the private sector.

A steady decline in support has generally been seen in the government's R&D funding policies. A brief from the Royal Society of Canada to the Prime Minister said that, although it was generally recognised in 1969 that government support for industrial R&D was already too small, it is now even smaller: in terms of the fraction of federal budgetary expenditures it had fallen by 1975 to 68% of the 1969 level.

Now a report has come from the MOSST* conveying the message that the Make-or-Buy Policy has in fact been rather successful, and that more of it is needed. The report is full of statistics supporting its argument, but its general tone appeared so contrary to what could be heard in non-government quarters that it left the impression of being a rather curious document.

A little investigation suggests why: the choice of statistics is highly selective, some say misleading. In fact it is questionable in some cases whether they show what they are purported to show. One table, for example, shows the increase since 1973-74 of mission-oriented contracts awarded to industry by the Department of Supply and Services on behalf of the various government departments, the intent being to indicate that they have increased. Yet the table lumps together all contracts from government departments, as though all resulted from the policy. In fact government departments have contracted out to industry for years, to a greater or lesser degree.

One diagram showing payments to Canadian industry for R&D between 1963-64 and 1975-76 shows not only that research contracts have increased since the Make-or-Buy policy was initiated, but also that, years before, in 1965-66, such contracts were larger than any year until 1973-74, indicating that the grants level fluctuates from year to year, and that taking just a short-term view (that is, from 1973-76) can be misleading. (The year 1965-66 was a good one largely as a result of the hydrofoil project of the Department of Defense.)

Analysis of the figures used in the report actually shows that government R&D contracts to industry had already begun to rise *before* the Make-or-Buy policy was put into effect, and, plotting them on a graph, that the level of R&D contracts to industry would be at the same point they reached even if they had merely been extrapolated at their former rate.

Furthermore, the ratio of payments to industry vs. intramural R&D was less in 1975-76 than a decade earlier by a substantial margin (even though that was only for a single year). Even since the Make-or-Buy policy was introduced, the increase in industrial contracts has amounted to only roughly $20 million, while the increase in in-house research funds has amounted to $57 million. And in terms of overall science expenditures by the federal government, industry's share has fallen from 19% to 16% during the three-year period of the policy.

The report, which was only released this summer, admits that it is premature to assess whether the ultimate economic effect aimed for in the policy is yet being achieved. But it does claim that the policy has had a beneficial effect on several sectors of industry. At the same time, it admits that during the course of the past five years, the proportion of Canada's total resources directed at R&D has steadily declined: gross national expenditure on research and development as a percentage of GNP had dropped in 1972 to 1.14% from 1.29% in 1969, while the Senate special committee on science policy had said it should be about 2.5% to keep Canada competitive internationally.

The report finds that the effects of the policy to date have been most pronounced in electronics, transportation and scientific services, where objectives of government and industry most closely coincide in the context of current departmental missions. If the programme were to be extended—which the report thinks advisable—either a reduction in government scientific staff or a substantial increase in funds for contracts would be necessary. □

Trudeau rings the changes

Stung by the results of a public opinion poll that showed Canada's Liberal Government to be lagging far behind the opposition in popularity, and by criticisms of both his bilingualism and anti-inflation programmes, the Canadian Prime Minister, Pierre Trudeau, recently announced wide-ranging Cabinet changes that have produced a new head for the Ministry of State for Science and Technology (MOSST).

Gone not just from the ministry, whose portfolio he held jointly with that of Public Works, but also from the government itself, is C. M. "Bud" Drury, who at 64 was one of the two most senior Cabinet members (with Mitchell Sharp, who has also resigned). In his place is Hugh Faulkner, a more junior minister who, according to the Toronto *Globe and Mail's* Ottawa correspondent, "was demoted from Secretary of State to Minister of State for Science and Technology mostly because the Liberal caucus felt that he had failed to sell bilingualism, a special responsibility of the Secretary of State."

Be that as it may, the new appointment is not calculated to cause rejoicing throughout the science community. Mr Faulkner's appointment makes it obvious that the science ministry will continue to be regarded by the government as a low priority, as it has been from its inception in 1971. Apart from Mr Drury, Mr Trudeau has always appointed inexperienced or junior ministers for this post.

The new appointment comes as just one more disappointment for the science community. But to many, Mr Drury's attitude towards the needs of science in Canada had in recent months seemed apathetic if not openly hostile. In spite of studies and recommendations on national science organisation, the government in general and Mr Trudeau in particular apparently lost interest in science and technology. The concerns of bilingualism and inflation had loomed so large as to eclipse science, and doubts had grown about the government's professed concern for the problems of science-based industry.

IN BRIEF

ESA moves

The European Space Agency's Scientific Programme Committee has approved the spending of about $115 million on two projects: a contribution to NASA's space telescope, due for launching in 1983, and the Geosari project involving the Geos spacecraft and Ariane launcher.

The $104 million provisionally allocated to the space telescope will provide 15% of the costs of the telescope and its associated Science Institute. In exchange European astronomers will get at least 15% of total observing time.

Negotiations with NASA on another ESA/NASA cooperative project, the Out-of-Ecliptic Mission, will continue. LIRTS (Large Infrared Telescope on Spacelab) and several other projects will be considered in the spring.

The council of the European Space Agency has appointed a new director of the Spacelab programme, M Michel Bignier. Until June 1976 M Bignier was Director General of the French Centre National D'Etudes Spatiales.

MRC report warning

The UK Medical Research Council (MRC), referring to the dual-support system for research in its annual report published this week, warns that, under extreme pressure, it would be contractually obliged to protect the security of its own 4,373 staff at its 68 research establishments, and the capital investment associated therein, if necessary at the expense of its indirect support for research at universities. Such a situation, the outgoing Secretary Sir John Gray emphasised, was not yet foreshadowed, and MRC policy remained one of parity between the two lines of support.

Government departments have redeployed with the MRC the 25% of the council's grant-in-aid from the Treasury transferred to them under the Rothschild arrangements. Total expenditure for the year amounted to £47.18 million, up 30%, of which international subscriptions, which grew 66%, were £960,000. Increases in these subscriptions because of exchange rate variations, along with pay and associated awards, made three additions to the grant-in-aid necessary. The current doctrine of cash limits, said Sir John, meant that any extra money to meet international obligations would have to come from ongoing work.

SGHWR inquiry continues

Following the summer recess, the UK House of Commons Select Committee for Science and Technology has resumed its inquiry into the decision surrounding Britain's involvement with the SGHW reactor with evidence from the Central Electricity Generating Board (CEGB) last week and from the United Kingdom Atomic Energy Authority (UKAEA) earlier this week.

In another development, a report from a study group chaired by Dr Walter Marshall on the integrity of pressure vessels for light water reactors was published by the UKAEA last week. The 156-page report makes 40 "essential recommendations" and has buttressed the case for PWRs and against SGHWRs in Britain.

THE pertinacity of the promoters of laetrile for cancer "cure" is probably related to the amount of funds at their disposal. Some possible clues are in a grand jury indictment in California, stating that 10 ml of laetrile, sufficient for three daily injections, cost American cancer patients as much as $60 and that a physician who administers laetrile "as a vitamin supplement" deposited $2.5 million in northern California bank accounts between May 1974 and August 1975. The president of the Committee for Freedom of Choice of Cancer Therapy was arrested last December with $40,000 worth of laetrile in his car, along with a loaded Browning automatic pistol. Previous cancer remedies such as Krebiozen and the Hoxsey remedy were on a smaller scale. The big question is: will political pressure, financed by profits from laetrile, overcome the opposition of organised medicine to its legalisation?

The Merck Index describes laetrile as the β-glucuronide rather than the β-glucoside, of L-mandelonitrile. It seems probable, however, that the laetrile of "underground commerce" is actually amygdalin (the β-glucoside). An argument used by the pushers of laetrile is that it is "vitamin B_{17}". In support of this, one of them stated that "it becomes almost impossible, on the negative side ever to declare scientifically that a compound is not a vitamin . . .". Constance Holden, a staff writer for *Science*, echoing this statement, said on September 10, that "no one has really been able to prove" that there is no such vitamin (as B_{17}). She also said that it is a "fact that scientists are not able to defini-

Laetrile struggles

THOMAS H. JUKES

tively refute the nutritional claims about laetrile . . .". Actually nutritionists feel no inhibitions about refuting such claims.

The cyanogenic glycosides, including amygdalin, have no resemblance to vitamins. The accepted definition of a vitamin is that it is necessary in the diet of at least one vertebrate organism to prevent a specific deficiency disease. The Committee on Nomenclature of the American Institute of Nutrition found no scientific evidence (i) for the existence of a nutrient identified as vitamin B_{17} or (ii) that laetrile has nutrient properties or nutritional value for "either animals or humans" (*sic*). Cyanogenic glycosides are often poisonous to livestock. One of them, linamarin, is present in cassava, and sometimes causes human blindness.

The persistent efforts of laetrile proponents were successful in Alaska where, on June 21, 1976, Governor Hammond allowed the "Laetrile Bill" to become law, thus permitting physicians to prescribe laetrile. His excuse was an implementation of the "equal time for nonsense" principle, which says that if there are two sides to an issue, it is not the role of a statesman to decide which one is right. The history of scientific medicine as applied to public health shows that it is insufficient to make discoveries that benefit human beings. It is equally important to get rid of quack remedies to make room for new advances. Until recently, legislators accepted informed advice in framing public health laws, but Alaska is "different".

The latest news is fascinating: the Assistant Attorney General of Alaska now says that sale or distribution of laetrile in Alaska by any person, including a physician, is illegal. There may be a problem in filling those prescriptions.

correspondence

Future of journals

SIR,—Your editorial of August 26 (p. 731) on the future of journals is typical of the current negative attitude to the problem. It is a pity that the more constructive contribution on the subject from Professor May and his wife (*Nature,* February 12, p. 446) was not mentioned.

The rising costs of running a library are eating into the money available for the purchase of publications, and yet we constantly read that there are too many journals, never too many librarians. The main library is often rightly the showpiece of the modern university campus, but are all librarians aware of their primary responsibility? Rows of expensively bound journals may satisfy the eye, but a drawer or binder of microfiche could be more relevant to academic needs.

The "pious hope" that two journals might be shut down for every new one launched would not be considered particularly pious by the users of the two guillotined journals. It would be more sensible in the short term to cut down on costs on both sides, amalgamate smaller, ailing periodicals and employ new methods such as those pioneered by the chemical societies in the United States and Britain.

We gather from various journal publishers that despite the economic conditions prevailing in many countries, reasonably priced research journals of medium circulations (1,000–4,000 subscribers) have generally enjoyed a steady increase in circulation, which has been maintained this year. This is an inevitable result of the worldwide growth of the academic community, and possibly also of the rationalisation of library purchases; for example, a university which may take 10 subscriptions to a major journal, or an 'essential title', for various departments, might cut down to two for the main libraries (*Nature* beware!), thus releasing funds for new acquisitions.

The long-term result of this trend might well be a huge number of journals distributed to, say, 1,000 resource/information centres scattered around the world. Researchers would then obtain photocopies of original papers, having browsed through synopses and abstracts, at their local library, by then reduced to little more than a post-office. This surely could lead to delays of a week or more. Faced

with such a possible outcome perhaps researchers and librarians might learn to love cabinets of microfiche.

May we therefore suggest that the present number of 35,000 journal titles should be viewed as a challenge, and not approached with despair. The final suggestion in your editorial, that authors should supply full details on request, would be a long step backwards. Imagine trying to trace an author from his last known address; with the present postal services it could take months. Who would be prepared to referee papers which might never be read? What use would data lists be in job applications? Evidence could be altered at a later date as no dated record would exist. Such chaos should be avoided. More organisation is required, not less.

> PETER ASHBY
> ROBERT CAMPBELL

Oxford Microfilm Publications Ltd.
Oxford OX1 4EY, UK

SIR,—I should like to comment on your editorial about the future of journals (August 26). As librarian in a research laboratory, I would agree with most of what you say—with the proviso that the changes you refer to should be reflected not only in journal subscription price, but in volume of paper. This is conspicuously not the case in the experiment of the American Chemical Society, where the weight of paper is much higher in the experimental than in the traditional journal.

As a practising organic chemist, I must, however, reply a firm "no" to your last question. Chemistry in general, and organic chemistry in particular, is an experimental science, and we must have the experimental detail readily available. One of the greatest weaknesses of the *Bulletin de la Société Chimique de France* was that experimental sections were abbreviated to the point of irreproducibility, although things have improved recently. There is a large body of opinion (perhaps even a majority) which recognises that theoretical sections are usually fitted to experimental sections after the event; and indeed, what some would like to see is a journal in which only the experimental section and the literature list is printed, the theoretical section being available on microfiche if the reader were unable to supply his own!

> ALAN F. THOMAS

Firmenich SA,
Geneva 8, Switzerland

Mistaken impression

SIR—I fear that Wil Lepkowski, in quoting from a paper that I presented at the Conference on Tradition and Change in Physics Graduate Education at Pennsylvania State University in August 1974, may have inadvertently given a mistaken impression of my position (August 12, page 528).

As I state in the same paper, "The Berkeley Group missed out on making these discoveries not so much in spite of the fact that their laboratory was so large, efficient, well-run and well-managed as precisely because of it".

The purpose of my paper was not to criticise a particular laboratory, but to argue, on the basis of historical experience, that the techniques of "Big Science", while they may be useful in solving certain engineering problems, as in the case of the Manhattan Project or the Apollo Program, are by their very nature not conducive to the making of fundamental, new scientific discoveries. Condensations of my paper were published in *New Scientist,* **63,** **462** (1974) and *Center Magazine,* 55 (May–June 1975).

> ROBERT J. YAES

Memorial University of
Newfoundland,
Canada A1C 5S7

Biological fly killer?

SIR,—The news on Human Trypanosomiasis Today by F. E. G. Cox (*Nature,* August 19, p. 646) calls for some views. As long as the immunological control of sleeping sickness remains unlikely, eradication of the genus *Glossinia* seems to be the only way to control the disease. Cox cited recent successes in Nigeria where a riverine area of 27,500 km² was sprayed aerially. The same technique was recently used over parts of Bangladesh in an attempt to control mosquitoes. The disappearance of quite a number of insect species was subsequently noticed, though all of them may well have made a come back, and for sure the mosquitoes did. Bearing in mind the delicacy of ecological balances nobody should be satisfied with aerial spraying as a method of controlling insect species. Intensive efforts should be mounted to introduce methods of biological control, such as releasing specific pathogens.

> K. M. S. AZIZ

Cholera Research Laboratory,
Dacca-2, Bangladesh

news and views

Non-histone proteins and transcription

from Carol K. Klukas

THE genomes of mouse and of humans, having molecular weights of 3×10^{12} daltons each, contain enough DNA to code for thousands of genes. Some of these sequences are present in multiple copy form while others are included only once; some are transcribed in all cells and their presence is essential for proper cell function while others code for proteins of more limited function and are transcribed only in certain specialised cells, or perhaps only at specific times in the development of an organism or even only at specific phases in the cell cycle; some sequences code for proteins which eventually must be produced in very large quantities in a cell while other sequences will dictate the synthesis of only a few protein molecules. How is it that all the sequences can be sorted out and be transcribed or not transcribed in just the right quantitative and temporal manner to result in proper function of a cell and of an organism? What factors are involved in the control of eukaryotic genome expression?

Since the answer to this complex question must lie at least in part in the composition and structure of chromatin (the DNA with the histones and non-histone proteins which are complexed with it *in vivo*), much current research is designed to answer various specific questions concerning chromatin. Three papers published back to back in the July issue of *Biochemistry* are noteworthy examples of such research.

In each of these papers, RNA transcription *in vitro* from isolated chromatin is compared on several criteria with *in vivo* transcription to determine whether *in vitro* transcription is a faithful representation of transcription as it occurs in the cell. Once this has been established, the composition of chromatin can be changed in a controlled way and the resultant changes in transcription can be analysed. This approach should reveal which aspects of chromatin composition and organisation are essential for the maintenance of *in vivo*-like transcription.

In the first of the papers, Bacheler and Smith (*Biochemistry*, **15**, 3281; 1976) compare total *in vitro* synthesised transcripts of mouse liver chromatin with nuclear RNA by means of RNA excess hybridisations with small amounts of radioactive mouse DNA which has been fractionated on hydroxyapatite into three classes each differing from the others in the reiteration frequency of the sequences it contains. The results of these experiments indicate that all sequences found in nuclear RNA are also transcribed *in vitro*. Furthermore, the transcription of reiterated sequences from isolated chromatin is restricted in just the same way as is transcription *in vivo*, that is, highly repetitious sequences in mouse DNA are not represented in either nuclear RNA or in RNA transcribed from chromatin *in vitro*. Competition experiments show that there are sequences transcribed *in vitro* from chromatin which are not present in nuclear RNA. But these belong to a very limited set of sequences unless the chromatin is extracted with 0.5 M NaCl before transcription. This salt extraction, which strips the chromatin of H1 histones as well as of certain non-histone chromosomal proteins, results in transcription which is no longer sequence restricted. Thus, Bacheler and Smith show that at least some significant aspects of *in vivo* transcription can be preserved *in vitro* and can be altered by salt extraction, suggesting the importance of H1 histone and/or of some of the acidic chromosomal proteins in transcriptional control.

Quite a different way of analysing *in vitro* transcripts, which is used in the other two papers (pages 3291 and 3296), is to assay for the presence of transcripts of one specific gene or set of genes, in this case the histone mRNAs. Stein, Reed and Stein have prepared ³H-cDNA complementary to histone mRNA and used it in hybridisation experiments to probe for histone mRNA in transcripts from various reconstituted DNA–chromatin complexes prepared *in vitro* from fraction-

ated HeLa cell chromatin. These studies show that naked DNA is an effective template for *in vitro* transcription of histone genes (see also Huang *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **48**, 1216; 1962; Allfrey *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **49**, 414; 1963) but that this transcription is inhibited nonspecifically by the addition of histones. Further, non-histone chromosomal proteins mixed with DNA in the *in vitro* system have no effect at all on the rate of histone gene transcription or on the frequency of histone mRNA in the transcripts unless however the non-histone proteins are mixed with the DNA in the presence of histones. In this last instance histone mRNA genes are made selectively transcribable, and thus histone mRNA represents a higher percentage of the total transcript. This result indicates that the interaction of histones and non-histone chromosomal proteins must be involved in a significant way in the control of specific gene transcription in this system and as the same effect of non-histone proteins has been observed in globin gene transcription (Paul *et al.*, *Cold Spring Harbor Symp. quant. Biol.*, **38**, 885; 1973; Barrett *et al.*, *Proc. natn. Acad. Sci. U.S.A.*, **71**, 5057; 1974; Chiu *et al.*, *J. biol. Chem.*, **250**, 9431; 1975) it is quite likely that histone–non-histone protein interaction is a general transcriptional control mechanism.

In these experiments the chromatin used was isolated from HeLa cells in the S phase of the cell cycle. In the last paper (page 3296) Park, Stein, and Stein investigate the *in vitro* transcription of G_1-phase chromatin and report that they find very little histone mRNA in the transcripts. But if S-phase chromosomal proteins are added to G_1-phase chromatin, the specific transcription of histone genes is enhanced. When G_1 and S-phase chromatins are fractionated and reconstituted using combinations of pooled DNA and pooled histones with varying amounts of G_1 or S-phase non-histone proteins, it becomes clear that S-phase chroma-

tin contains a non-histone protein which has the ability to render the histone genes available for transcription. It is the absence of the protein in G_1-phase chromatin rather than the presence of some inhibitor molecule which distinguishes G_1-phase from S-phase chromatin. Thus the importance of the non-histone proteins and their interaction with DNA and histones in the control of specific genes is again indicated.

The type of *in vitro* transcription and reconstitution experiments presented in these papers are beset with difficulties not only in their execution, but also in their interpretation. Often such work meets with much criticism and scepticism because factors such as just how the DNA probe was made and how the reconstitution and hybridisations were performed can alter the results significantly and consequently lead to misinterpretations. Also one can always criticise this type of work by pointing out that the studies are after all *in vitro* and may have nothing to do with the transcriptional process which occurs in the cell. Certainly this is correct; however, *in vitro* transcription and reconstitution experiments despite their difficulties are useful in pointing out which components of the *in vivo* system may be of particular significance. In a system as complex as that for the control of eukaryotic genome expression, such clues to which factors are likely to be most important is certainly useful as a guide for further thought and research. ☐

Higher ozone concentrations over Britain

from Peter D. Moore

THE elevated levels of ozone in the atmosphere of London and other parts of Britain have been a source of concern for some years (see *Nature*, **256**, 537; 1975). Man-made pollution is undoubtedly a major source of ozone in the troposphere and inflates the 2–4 p.p.hm. which is normally found there. Petrol and diesel engines are considered to be important in this process of ozone generation, for they emit nitric oxide and also uncombusted hydrocarbons which oxidize the nitric oxide to nitrogen dioxide. NO_2 is liable to photochemical dissociation in daylight leading to the generation of atomic oxygen which interacts with molecular oxygen to produce ozone, O_3.

Ozone is a powerful oxidant and is known to damage both animal and plant tissues (Bell and Cox, *Environ. Pollut.*, **8**, 163; 1975). The Environmental Protection Agency of the United States has established an air quality standard of a maximum ozone concentration of 8 p.p.hm. for an hour's duration. In London, the Greater London Council's guideline has been set at the same level.

For some years this guideline has been exceeded, usually during anticyclonic periods in summer. On July 13, 1972, a peak value of 11 p.p.hm. O_3 was recorded in Central London (Derwent and Stewart, *Nature*, **241**, 342; 1973). On page 580 of this issue of *Nature* Ball presents data for the summer of 1975 in London and registers a peak value of 15 p.p.hm. on June 26, 1975, almost twice the GLC guideline level. This is the kind of concentration at which, according to the WHO and the US Department of Health, a degree of eye irritation and aggravation of respiratory diseases may be expected.

This is still well below the 28 p.p.hm. of Sydney, over 30 p.p.hm. in Tokyo and up to 99 p.p.hm. in Los Angeles. It is, however, a disturbing trend. Ball has kindly provided some preliminary figures for the current (1976) summer in London which suggest that the trend towards higher ozone concentrations in Britain is continuing. Even in early May of this year, levels of 18 p.p.hm. had been recorded in London. Subsequently values up to 21 p.p.hm. have occurred, and the guideline concentration has often been exceeded for as many as 40 h per week over the whole of London. Records from Harwell are even higher than these.

Ball's data do help to resolve the question of whether Britain's ozone is largely derived from the Continent or whether our cities make a significant contribution. Cox *et al.* (*Nature*, **255**, 118; 1975) showed that photochemically generated ozone can be transported distances of the order of 100–1,000 km. Continental ozone can enter the British troposphere in sufficient concentration to raise ambient levels to as much as 15 p.p.hm., as was shown during easterly air movements at a recording station in Suffolk during August 1973. Stewart *et al.* (also in this issue of *Nature*, page 582) show that high ozone concentrations occur most often on days when the wind direction is north-east, through to south This lends some support to the contention of Cox *et al.*, but it is also true that when anticyclonic conditions prevail the wind is most often in that quarter. As Stewart shows, ozone is highest when

temperature and insolation are high, and these conditions are also associated with anticyclones. Ball now demonstrates that among the London recording stations higher values of ozone are found at those sites which receive air masses from over Greater London, and concludes that home-grown ozone represents a significant contribution to our tropospheric load.

To what extent are these recent high levels of ozone a consequence of the unusually sunny, anticyclonic summers we have experienced of late? Only continued monitoring of gaseous pollution over London can answer this. It is salutary to note that in cities like Sydney, ozone became a problem with startling suddenness. A further, unanswered question is the degree to which ozone interacts with sulphur dioxide, our other main gaseous pollutant, in its harmful effects. Menser and Heggestad (*Science*, **153**, 424; 1966) showed that tobacco plants were severely injured by mixtures of SO_2 and O_3 in concentrations of 24 p.p.hm. and 3 p.p.hm. respectively, whereas neither of these concentrations produced damage on their own. It is possible that this synergistic effect is general, in which case ozone may be an even more serious pollutant than is currently predicted, especially in cities such as London where SO_2 concentrations are still high. ☐

Pleistocene mammals

from D. W. Yalden

Historically, the evidence provided by fossil mammals played a large part in establishing the glacial theory of Pleistocene geology. The discovery in the 1860s of fossil lemmings in southern England, in particular, was readily accepted as evidence of former tundra conditions. Though early excavations of caves and river terraces concentrated on large mammals, such as mammoths, or on human remains and artefacts, much information on smaller mammals was obtained; this formed the basis as long ago as 1926 for a large (though uncompleted) monograph by Hinton (*Monograph of Voles and Lemmings (Microtinae) Living and Extinct*, British Museum, (Natural History), London). While this early work established Quaternary biology as a science, it suffered, as pioneering work inevitably must, from being too early. The fossil mammals were subjected to an extremely typological classification, with names founded on extreme variants, and more normal intermediates left unnamed. The very complex nature of

ALTHOUGH it was the first nuclear process studied by Rutherford, alpha decay is still not properly understood. A large number of alpha decay rates have now been measured for heavy nuclei, and these depend basically on two factors, a nuclear structure factor that gives the probability that there is an alpha particle in the nuclear surface 'waiting to get out', and a penetration factor that gives the probability that it will be able to tunnel quantum-mechanically through the barrier and so escape from the nucleus.

Ever since Gamow first applied quantum mechanics to the nucleus, it has been known that the penetration probability depends very critically on the height and width of the potential barrier through which the alpha particle has to tunnel, and hence on the energy of the emitted alpha particle. This is why the alpha decay halflives vary from a very small fraction of a second to many thousands of years, while the corresponding alpha energies vary by less than an order of magnitude.

This simple theory was very successful in accounting for the relative alpha decay rates of a large number of radioactive nuclei, but the absolute values were much less certain, due to the lack of knowledge of the nuclear structure and of the potential barrier. Using the best available nuclear structure information, there was usually a factor of a thousand between the measured and calculated decay rates and it was generally considered that this large factor was still within the uncertainties of the potentials.

This point has now been carefully checked by De Vries at Rochester and Lilley and Franey at Minnesota (*Phys. Rev. Lett.*, **37**, 481; 1976). They

Absolute alpha decay rates

from P. E. Hodgson

studied the interaction of alpha particles with ^{208}Pb and ^{209}Bi and fitted the experimental data for elastic scattering at 22 MeV and the reaction cross section from 17 to 24 MeV by a series of optical model potentials. The reaction cross sections are known accurately since they are the sum of the (α,n), $(\alpha,2n)$ and (α,γ) cross sections, all of which can be measured accurately. No other reactions are energetically possible.

They found that this data imposes tight restrictions on the optical potentials, so that an acceptable fit can only be obtained with parameters constrained within narrow limits. The penetrabilities calculated from these potentials vary somewhat, but far less than was previously thought. This makes it possible to calculate accurate absolute values of the quantity called the reduced alpha width γ^2 that contain the nuclear structure information alone. It is found that these values are still about a thousand times greater than those calculated from the shell model. For example, for ^{210}Po, the value of γ^2_α found from the measured alpha decay rate and the calculated penetrability is 0.88 $^{+4.4}_{-0.19}$, whereas the value calculated from the shell model by Harada is only $(0.31$ to $3.3) \times 10^{-3}$. Results for other nuclei are similar.

This is a serious discrepancy and shows the need for much more sophisticated nuclear structure calculations, including more configuration mixing and the renormalisation required by antisymmetrisation. A thousand times more alpha particles are emitted than would be expected from existing shell model calculations, so it may well be that this is due to substantial clustering of alpha particles on the nuclear surface that cannot yet be described by the shell model. ☐

---

the climatic fluctuations was not appreciated either (indeed Hinton was a monoglacialist, and so assigned faunas either to "the" cold period or to the preceeding warm period), with the result that much ecological and stratigraphical information was overlooked or lost.

Biological aspects of Pleistocene study have been revolutionised first by pollen analysis and more recently by the study of insect, especially beetle, faunas; physical aspects of study have added absolute chronologies (from radiocarbon and other dating techniques) and absolute temperature calibration (from oxygen isotope ratios). In this surge of knowledge, Pleistocene mammals have, at least in Britain, been left far behind; Zeuner's book (*The Pleistocene Period*, Hutchinson, 1944) contained a chapter of 30 pages on mammals, but only four single-line entries on pollen analysis, whereas West (*Pleistocene Geology and Biology*, Longman, 1968), has 32 pages on pollen stratigraphy but only six on mammals.

Clearly the time was ripe for restudy of British Pleistocene mammals, not only in their own right but also to integrate them into the newer framework of Quaternary chronology. Some significant signs of such restudy are now available. One is Stuart's review relating all available vertebrate occurrences to the pollen stratigraphy of Britain (*Biol. Rev.*, **49**, 225–266; 1974). this involved new collecting from sites where information on pollen was available or obtainable as well as a re-examination of a scattered and often obsolete literature. Another equally valuable review of a different type is Maglio's paper on the evolution of the elephants (*Trans. Am. phil. Soc.*, **63**, 1–149; 1973). Because they are so conspicuous, fossil elephants, or at least their molars, have played an important part in stratigraphy, but their taxonomy has been extremely muddled.

Perhaps the most important such contribution has just been published by Sutcliffe and Kowalski (*Pleistocene Rodents of the British Isles, Bull Br. Mus. (nat. Hist.) Geol.* **27**, 31–147; 1976). This account is important in a number of respects. For a start, the rodents are, and have been throughout the Pleistocene, the most numerous mammals; they are therefore the most abundant in cave deposits, and so potentially the most useful archaeologically or stratigraphically, and they are arguably the most important ecologically. Their potential has however been largely unrealised, partly perhaps because archaeologists have tended to ignore them, but largely because, as a result of the early studies, they were in a taxonomically unsatisfactory state and therefore unreliable stratigraphically. This review should at least resolve these last obstructions. The literature records are thoroughly reviewed, and their stratigraphical assignments reassessed. The taxonomy is also reviewed, so that it is now much clearer just how many distinct species were present in the Pleistocene. The dubious taxa erected by Hinton, among others, are thoroughly scrutinised, and many are relegated to synonymy. Of particular importance is the cautious reassessment of *Microtus* species, especially *M. arvalis*. This vole has a rather southern distribution in Europe, being absent from Scandinavia and, except for Orkney and Guernsey, from Britain. The occurrence of fossil remains attributed to it, however (often as *M. corneri*) have led to suggestions that it was an early postglacial immigrant, since isolated as a relict in Orkney. Not only is this inherently unlikely on ecological grounds (human introduction being more likely an explanation) but, as this review makes clear, the supposed fossil occurrences are very doubtful indeed. The

review also reveals however many very real problems which require further study or perhaps further material. For example, we do not know for certain when lemmings disappeared from Britain nor whether, as the evidence suggests, the high Arctic collared lemming, *Dicrostonyx torquatus*, survived later than the more southerly *Lemmus lemmus*.

More important, however, than this sort of query which the review poses is the framework which it provides for evolutionary studies. G. R. Coope (personal communication) has averred that beetles, because of their mobility, provide a better indicator of past climates than plants, but that their reliability for this depends on their evident morphological stability and presumed ecological stability throughout the Pleistocene. Plants also seem to have evolved little. The rodents, however, share the mobility of beetles, but at the same time demonstrate considerable evolutionary flexibility. Their rates of evolution and the relationship of these to ecological changes are aspects which should begin to attract attention.

Cosmic rays at Leeds

from a Correspondent

> The fifth European Cosmic Ray Symposium was held at Leeds on September 14–17, 1976.

COSMIC ray studies still have a lot to offer the nuclear physicist. That at the very least was clear from the discussions at the symposium. The energies available with cosmic ray particles (up to $\gtrsim 10^{20}$ eV) still far exceed the best accelerators. Thus effects such as high transverse momentum and increasing cross sections for hadron interactions that are perhaps beginning to appear at the highest machine energies should be clearly observable from cosmic ray studies. If, as accelerator work initially seemed to indicate, Feynman scaling theory holds then it should become very evident at these high energies.

Experimentally, problems arise in studying high energy cosmic rays because of their inacessibility and very low flux. The high energy primary particles reaching the Earth interact in the atmosphere causing a cascade of secondary particles to reach sea-level in the form of an Extensive Air Shower (EAS). Since the low primary flux makes direct collection virtually impossible these EASs are the only convenient means of investigating the nature and properties of the high energy particles.

A major topic at the Leeds symposium was the attempt to relate the experimentally measured properties of the EAS to theoretical predictions based on cascade calculations assuming both a nuclear physics model and a particular mass composition spectrum for the primary particles. High transverse momentum phenomena in nuclear interactions (diverging from average CKP behaviour) have been a topic of discussion at cosmic ray conferences for many years, and now positive evidence is also appearing from accelerators. Controversy still reigns, however, as to the magnitude of the effect or indeed whether the observed events could be just the expected high momentum tail from the CKP distribution.

The scaling theory of nuclear interactions came in for a lot of hammering. Evidence is accumulating from EAS research against 'scaling' being operative at high energies. Here a problem in interpretation is the uncertainty in the mass composition of the primary cosmic ray particle initiating the EAS. Assuming the established composition at low energies (~80% protons) then 'scaling' appears to be incompatible with experimental EAS data. However if predominantly iron primaries are assumed to occur ~10^{17} eV the 'scaling' theory can perhaps be rescued. An additional concern, pointed out at the symposium, was the as yet inadequate treatment in the model calculations of extreme fluctuation effects coupled with the steeply falling primary energy spectrum. Rigorous calculations may well make 'scaling' more acceptable. In passing it should be noted that recent results from Fermilab and Stanford on deep inelastic scattering also show departures from 'scaling'.

The most exciting experimental work in EAS studies discussed at the symposium was reported by K. E. Turver and colleagues from the University of Durham. This group have very carefully measured the Cerenkov radiation produced in the atmosphere by EAS. These observations were recorded at the Haverah Park detector array, North Yorkshire, from EAS of primary energy >10^{17} eV. Although such atmospheric Cerenkov radiation has been investigated previously the advance by Turver's group was achieved by making use of a well-spaced array of eight photomultiplier detectors and correlating their responses. By analysing the pulse shape from the different detectors the researchers claim to be able to pinpoint the EAS development with a considerable degree of accuracy. It is fairly early days in their investigations but the technique looks very promising. The main problem at the North Yorkshire moors array is the available operation time as clear dark

nights are required—only 100 hours was achieved last winter. However, if this technique is established at Haverah Park, in close collaboration with the particle detector array there, then it should be possible to set up such an air Cerenkov array independently in a more favourable climate.

The origin and mass composition of the high energy primary cosmic rays still remain the key astrophysical questions to be answered. The extragalactic *versus* galactic origin controversy continues with almost equal support on both sides. Until recently it seemed that cosmic rays even at the highest energies appeared to reach the Earth equally from all directions. However, as the statistics of the really high energy events ($\gtrsim 10^{18}$ eV) built up in recent years there does appear to be some evidence for a small degree of anisotropy in right ascension. A. M. T. Pollock and A. A. Watson (University of Leeds) presented data on the arrival direction of primary particles of energy 10^{17}–10^{18} eV. Some 45,000 EASs were subjected to harmonic analysis in right ascension. A significant amplitude (0.1% chance probability) occurs in the direction of $251(\pm 15)^\circ$ RA. Evidence from the same group at higher energies (>10^{18} eV) is not so conclusive. There seems to be a change of phase with increasing energy but it is regarded as significant that the two highest energy events ($\gtrsim 10^{20}$ eV) come from close to the North Galactic Pole. D. D. Krasilnikov (Institute of Cosmophysical Research, Yakutsk, USSR) has plotted the arrival directions of the total world accumulation of 67 reported events greater than 4×10^{19} eV. An equatorial region in the galactic coordinate system seems to be practically devoid of events. Krasilnikov concludes that this can be interpreted as the result of screening by magnetic fields of an isotropic flux of the particles coming from outside the Galaxy radiodisk. If these particles were of galactic origin they would be required to be heavy nuclei from a powerful source in the centre or behind the centre of the Galaxy and influenced by a highly extensive radio halo with a very specific regular magnetic field structure. It is more likely that they are of extragalactic origin. Attempts to correlate arrival directions with pulsars have proved negative.

This symposium represented a special occasion in that for the first time the lower energy cosmic ray work (modulation section) was discussed at the same meeting as the high energy and EAS work. Usually two separate symposia are held at different venues. The amalgamation was in honour of Professor J. G. Wilson whose life-long interest in cosmic rays has spanned both branches of the research field. John Wilson is this year retiring from

the Cavendish Chair of Physics at the University of Leeds which he has held since 1952. His research in cosmic rays started under C. T. R. Wilson and he became a prominent member of P. M. S. Blackett's group at Manchester after the Second World War. Latterly he has been responsible for setting up and developing the EAS detector array at Haverah Park. He plans to continue his interest in the field during his retirement.

J. G. Wilson presented a paper dealing with the energy spectrum of cosmic rays at energies $\gtrsim 10^{17}$ eV. Evidence continues to mount in support of a flattening of the spectrum above 10^{19} eV. There appears to be no cut-off at this energy arising, as predicted, from interaction with the supposed universal 3 K photon flux. This inconsistency still remains a problem as the arrival direction analysis does seem to suggest an extragalactic origin for these very high energy cosmic rays.

In general then cosmic ray research still appears to be a hive of activity with plenty of promise. However, progress is inevitably slow, perhaps reflecting the difficulties of both technique and interpretation. ☐

Efficient bumblebees are almost specialists

from John Krebs

THE familiar sight of bumblebees gathering pollen and nectar from wild flowers is the basis of a recent study by B. Heinrich (*Ecol. Monogr.*, **46**, 105; 1976), in which he analysed the foraging behaviour of individually marked worker bees of the genus *Bombus*. The bee uses different techniques to collect pollen and nectar from different plant species, for example it collects pollen from dogwood by pressing its body on the flat inflorescence, while on the cranberry it hangs upside down and shakes the pollen onto its body. The bees in Heinrich's study area preferred certain flower species over others: *Chelone* being the most favoured and *Solidago* the least. In fact the rank order of preference (measured by the number of bees per flower) corresponded exactly to the rank order of potential nectar yield from the various species—I will return to this point later. While this picture applies to the whole bee population, each individual bee specialises on a particular plant species (its 'major') and occasionally visits one or two other species ('minors'): the overall distribution of foraging effort by the bees is the sum of many specialists on different plants.

Specialisation is not purely a result of attachment to one particular site, nor is it temporary; many workers probably retain the same 'major' throughout their life of four to six weeks. The 'minors' of individual bees are also constant, but Heinrich found that he could induce workers to adjust their preference by adding sugar droplets to a 'minor' in their repertoire, which then became the 'major', showing that the choice of a 'major' is related to how profitable different plants are. The benefit to a bee of specialising was revealed in an experiment in which Heinrich compared the efficiency of experienced and naive bees in extracting nectar from monkshood, which is unusual in having its nectar hidden in two modified petals. The naive bees, which had foraged on other flowers but not monkshood, had difficulty in finding and extracting the nectar while the specialists did not. This and other observations suggested that a non-specialist bee would be rather unsuccessful in competing for nectar and pollen with a specialist.

Heinrich's results pose two intriguing questions: Why do all bees not major on the most profitable flower species? And why do individuals maintain minors in their repertoire? The answer to the first question is straightforward. As I have mentioned, the bee population as a whole distributed its foraging effort according to the order of profitability of the various flowers, so that the commonest major is the top ranking nectar source, the next commonest major the second ranking plant and so on. It also turns out that after the foragers have exploited an area, all the different plants contain about the same amount of left-over nectar—they are all reduced to the same level of profitability. This means that the lower nectar content of poor quality flowers is (more or less) exactly compensated for by the fact that more bees compete for nectar in the rich flowers. (This pattern of harvesting, which has also been observed in birds, is predicted by theoretical models describing how a maximally efficient forager ought to behave, but more of that another time.)

The question of why bees maintain minors is more difficult, but G. Oster and B. Heinrich (*Ecol. Monogr.*, **46**, 129; 1976) have developed a model to offer a tentative explanation. Their argument goes like this: if there is an array of flowers giving different quantities of nectar per unit effort, an efficient worker with perfect knowledge should forage purely in the most profitable flowers. When the bee has no previous knowledge, it has to sample the flowers to estimate their profitabilities and then adopt the "pure strategy" of foraging in the best

flowers. But suppose Nature is awkward, and the different flower types continually change ranks, what should the bee do? In the extreme case, when Nature obeys Murphy's Law and what the bee thinks is best immediately becomes the worst, a "mixed strategy" is best because it insures against a sudden change. Assuming that Nature is perverse but not too perverse, the bee's best strategy lies somewhere between mixed and pure: majoring with one or two minors. ☐

Photobiology

from R. P. F. Gregory

> The Seventh International Congress on Photobiology was held in Rome on August 25–September 3, 1976.

WHAT advances stand out after four years' study of photobiology? G. Porter (Royal Institution, London) gave one answer in terms of picosecond kinetics. The primary processes of light-absorption by a pigment and its re-emission in fluorescence can now be studied at times approaching a few picoseconds using the streak camera. (The absolute theoretical barrier was set by the Uncertainty Principle in the region of 0.5 ps.) Both by advances in the understanding of the behaviour of excited states of pigments and the application of ultrashort pulsed laser spectroscopy, models have been produced of the operation at the nanosecond timescale of three photobiological pigments. M. Ottolenghi (Hebrew University, Jerusalem) described the interconversions of rhodopsin, bathorhodopsin and isorhodopsin in terms of the energy level diagram and was able to account for the observed quantum yield. Remarkably similar behaviour was reported for the analogous pigment bacteriorhodopsin from the typical phototroph *Halobacterium* (R. H. Lozier and colleagues, Ames Research Center, California). The phototransformations of phytochrome, preceding the appearance of the Pr and Pfr forms, are also in the model stage (R. E. Kendrick, University of Newcastle-upon-Tyne). Outside the formal sessions it was apparent that the problems of pigment organisation and the spectroscopic methods presently available for their investigation were common to these topics and also to the study of chlorophyll, both in membranes and in pigment-protein complexes. At this level Photobiology very clearly exists as a discipline.

The importance of flavins or flavoproteins in many responses of organisms to light is becoming apparent, and the much debated alternative, carotene, is on the retreat, in such examples as the flagellar response of *Euglena*, and the development of the coleoptile of oat seedlings. Flavoproteins also appear to provide an explanation for the various 'blue-light effects' in mitochondria, *Neurospora* and the alga *Protosiphon botryoides*. In several cases there is clear evidence that the flavin is coupled to an electron transport system, such as a *b*-cytochrome or plastocyanin. In these examples the metabolic links between the action of the pigments and the observed effects on the organisms were relatively obscure; however, in photosynthesis there have been impressive developments in metabolism. These cover both the chloroplast–cytoplasm relationship of the normal (C3) system and the agriculturally-wasteful process of photorespiration. H. R. Bolhar-Nordenkampf (The University, Vienna) described techniques of measuring photosynthesis and photorespiration in crop leaves, from which the behaviour and future yield of the crop could be assessed. The possibility of photorespiration being an essential detoxification system for the various oxide radicals continues to be discussed.

The subject very clearly divides between the above 'natural, positive' aspects, and the adventitious and negative effects of light, particularly ultraviolet. Although there is considerable data on the effects of ultraviolet on proteins and nucleic acids, the carcinogenic action on human skin could not be simply explained. A more satisfactory molecular explanation could be claimed concerning the success in treating certain skin diseases such as psoriasis and mycosis fungoides with psoralens (which bind to DNA) and UV-A light (M. A. Pathak, Harvard Medical School). Not only ultraviolet, but visible light was clearly shown by J. Marshall (Institute of Ophthalmology, University of London) to lead to striking degradation of retinal cones in animals and people subjected to artificially extended daylength.

The contribution of photobiology to economic problems occupied much attention, in that in principle solar energy remains the one large source of energy available for immediate use. Biological conversion was discussed by D. O. Hall (King's College, London) who pointed out that in many areas existing knowledge and technology could provide not only increased food crops but materials such as ethanol and glycerol at competitive prices. In several cases it had been shown that producing crops under cover in large areas with varied CO_2 levels not only relieves losses by photorespiration, but also increases nitrogen fixation and water economy. W. J. Oswald (University of California, Berkeley) is of course unchallengeable in his presentation of algal sewage-farming at a profit. A. A. Krasnovsky (Academy of Sciences, Moscow) and others set out possible photoconversion systems based on isolated pigments or stabilised membrane systems, provided with enzyme systems, for the production of hydrogen. It remained doubtful, however, whether simulated photosynthesis for the production of, essentially, heat would be competitive with existing 'solar panels' for heating (and refrigerating). The discussion of this topic (as Porter had the opportunity to point out in the Farrington-Daniel memorial lecture) has had a phenomenal growth rate in the past four years, and raised Photobiology to the status of a political issue. □

Sicily as the sum of parts

from Peter J. Smith

THE notion that Sicily straddles lithospheric plates is by now a familiar one. The crystalline massif forming the extreme northeast of the island is a continuation of the Calabrian metamorphic belt and thus linked geologically with Europe. The Ragusa carbonate platform in the extreme southeast, on the other hand, is an extension of the Sahara platform and thus linked geologically with Africa. The boundary between the African and European plates (see McKenzie, *Geophys. J.*, **30**, 109; 1972) probably runs along the northern edge of Sicily, except in the east where it lies south of the massif.

This interpretation places the re-sedimentation basin and nappes which form the western and central parts of the island firmly on the African plate. Indeed, Barberi *et al.* (*Earth planet. Sci. Lett.*, **22**, 123; 1974) were apparently able to confirm this when they showed palaeomagnetically that, the crystalline northeast apart, Sicily has been part of the African plate since the Cretaceous. But as might be expected from Sicily's marginal position, things may not be that simple. Caire (in *Gravity and Tectonics*, Wiley, 1973), for example, has suggested that there may have been a second intra-island plate boundary isolating the Ragusa platform—a possibility apparently not excluded by the work of Barberi and his colleagues who analysed rocks only from the extreme east.

To test this idea, Schult (*Earth planet. Sci. Lett.*, **31**, 454; 1976) has now carried out a palaeomagnetic study of Upper Cretaceous and Middle Jurassic volcanic rocks from the west of Sicily. The two poles thus obtained are not only significantly different from those of comparable ages for Europe (as expected) but are even further removed from those for Africa and the Ragusa platform. Since the Upper Cretaceous, western–central Sicily has apparently rotated clockwise by about 90° with respect to southeastern Sicily (and hence with respect to Africa) and by about 60° with respect to stable Europe.

Schult admits that the margin of error on his results is larger than one would wish; but taken at face value the new poles confirm Caire's proposal that western–central Sicily and the Ragusa platform once behaved as, or as part of, separate crustal blocks. And implicit in that confirmation is western–central Sicily's unique sense of rotation. All other known crustal block rotations in the central Mediterranean were apparently anti-clockwise. □

articles

Stratospheric aerosols and climatic change

Betty Baldwin, James B. Pollack, Audrey Summers & Owen B. Toon

Space Science Division, NASA–Ames Research Center, Moffett Field, California 94035

Carl Sagan

Laboratory for Planetary Studies, Cornell University, Ithaca, New York 14853

Warren Van Camp

Informatics, Inc., Palo Alto, California

Global heat balance calculations imply that stratospheric aerosols, generated by volcanic explosions, have made important contributions to some observed climatic changes but that such aerosols generated by supersonic transports and Space Shuttles are unlikely in the next few decades to have a significant climatic impact.

THE stratosphere contains a layer of submicron-sized particles, centred at an altitude of $\sim$ 20 km and extending from the tropopause to an altitude of $\sim$ 30 km (ref. 1). Usually, the particles of this layer are made primarily of a concentrated sulphuric acid solution[2,3], which is generated *in situ* from sulphur-containing gases that undergo a series of oxidation and hydration reactions[1]. At present, volcanic explosions are very important sources of these aerosols because of the sulphurous gases they inject into the stratosphere[4]. Volcanic explosions also introduce silicate particles into the stratosphere[4,5]. Because the gas–particle conversion has a time constant $\sim$ 1 yr (ref. 1), silicates are the dominant aerosol species in the first few months following a volcanic explosion[5], but H_2SO_4 dominates for the next few years[4]. A future source of stratospheric aerosols may be from flight through the stratosphere. The effluent from supersonic transports (SSTs) includes sulphur-containing gases, which will be converted to H_2SO_4, while the exhaust from Space Shuttles contain tiny aluminium oxide particles[1,6,7].

It is important to understand the climatic changes that might be produced by stratospheric aerosols. One variable useful in characterising the climate is the globally-averaged surface temperature, T, which depends on the global heat balance between the amount of solar energy absorbed by the entire Earth and the amount of thermal radiation emitted to space. Stratospheric aerosols scatter and absorb solar radiation and interact with and produce thermal radiation. Changes in their number will thus cause a temporary imbalance between the solar energy absorbed and the thermal radiation emitted to space and necessitate an adjustment in T to restore the global heat balance. We perform here accurate radiative transfer calculations to estimate the sensitivity of the climate to changes in the number of stratospheric aerosols. The results, when combined with observed changes in the level of volcanic activity, permit us to assess volcanism as a cause for past climatic changes. Calibrating by these findings, we calculate the climatic impact of future vehicular traffic through the stratosphere.

The model

We here summarise our radiative transfer calculations of T, specify the sources of our model's parameters, discuss factors not included in our calculations, and, where possible, introduce approximate corrections for these factors. It is not possible to evaluate the full climatic impact of changes in the number of stratospheric aerosols; but because the direct effect of these particles is radiative, our calculations provide a useful measure of the sensitivity of T to such changes.

For a specified model atmosphere, including a vertical profile of the aerosols, we determine the value of T that leads to global heat balance. Climatic change is measured by ΔT, the difference in T between the perturbed and unperturbed state. An accurate method, based on the doubling principle[8,9], was used to solve the multiple scattering problem for unpolarised light and so determine the amount of solar energy absorbed by the Earth. The thermal radiation emitted to space was found by numerical evaluation of the formal solution of the equation of radiative transfer[10]. For the solar calculations, account was taken of the absorption and scattering from both gases and aerosols. Because the single scattering albedo of stratospheric aerosols is small in the thermal infrared[11], scattering was neglected in the thermal calculations, but the opacity of both gases and aerosols was included. While no allowance was made for diurnal, latitudinal or seasonal variations in the model parameters, we did include vertical inhomogeneities. Also, tropospheric water clouds are present over 50% of the globe. A more detailed discussion is given elsewhere[4,7,8].

An increase in the amount of stratospheric aerosol can lead to either a warming or cooling of the Earth. Our radiative transfer calculations, as well as those of others[12], indicate that the sign and magnitude of ΔT depends critically on the real and imaginary parts of the refractive indices of the added particles, their size distribution, and the surface albedo. We use the measured refractive indices for materials constituting the aerosols of interest[13-15] and size distributions based upon available measurements[11]. Finally, we use a value of 0.1 for the surface albedo[16].

The increase of the optical depth of the stratosphere, $\Delta \tau$, caused by the addition of particles has a strong effect on the magnitude but not the sign of ΔT. This parameter is proportional to the number of added particles which are large enough to interact efficiently with light. Figure 1 illustrates the variations of $\Delta \tau$ over the past century, as determined from observations of the amount of unscattered sunlight and starlight transmitted by the atmosphere[4], and many of the observed
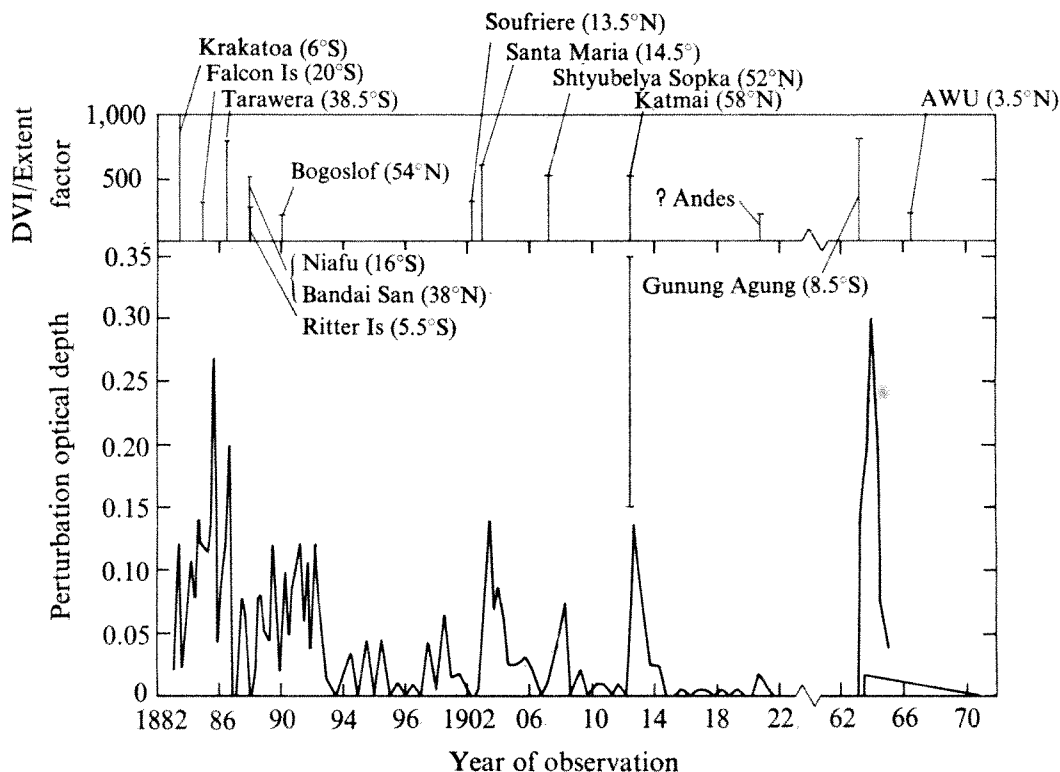
Fig. 1 Optical depth perturbations of the atmosphere at a wavelength of 0.55 μm as a function of time, and volcanic explosions that Lamb[25] has judged to have injected significant quantities of dust into the stratosphere. The dust veil index (DVI) is Lamb's assessment of the importance of an eruption. The larger perturbation in 1963 is a Southern Hemisphere value, and the smaller a Northern Hemisphere value. The error bar in 1912 represents a measurement at one location of a single day's perturbation and a one-month average. The solid curve is an average of monthly mean optical depths at several locations. Note the phase lag between the times of eruption and peak optical depths—partly because of the slow transport of dust from the point of injection into the stratosphere and partly because of the slow photochemical conversion of sulphur-containing gases to sulphuric acid aerosols, as is observed directly[37].

changes in $\Delta\tau$ can be associated with major volcanic events. We estimated $\Delta\tau$ for other times from less direct information, as indicated in the footnotes to Tables 1 and 2. In making these latter estimates, we have neglected any variation in the stratospheric residence times of the aerosols, which could result from changes in the stratospheric temperatures.

Our calculations neglect several feedback mechanisms that might augment the temperature changes (alterations of the snow and ice cover), retard these changes (an alteration in the dynamical heat transport), or has either effect (variation in cloudiness). Current climate models that include the first two of these feedback mechanisms suggest they cause an amplification by a factor of 4 in ΔT for transitions from interglacial to glacial conditions[17,18]. Since 10^3–10^4 yr are required to establish the great continental ice sheets[19,20], we multiply our results by this factor when dealing with changes on time scales $> 10^3$ yr, and use a smaller factor for time scales as short as a few hundred years[21].

Our calculations are also limited to the steady state, in which perturbation persists for a much longer period, P, than the response time of the land–atmosphere–ocean system, t. When dealing with time scales of a few months or less, when $P < t$ for both the troposphere and ocean[4], we hold the tropospheric and ground temperatures constant and consider only temperature changes in the stratosphere. For somewhat longer time scales (still < 4 yr) $P > t$ for the troposphere, but $< t$ for the mixed layer of the oceans[4]. In this case we multiply[4] our calculated value of ΔT by $P/2t$.

Volcanic activity as an agent for climatic change

Using this model, we have calculated ΔT for past, observed changes in the level of explosive volcanism and have compared the results with contemporaneous observations of ΔT. Table 1

summarises the results. The first two rows refer to the effects of single volcanic explosions, while the remaining rows refer to the effects of multiple volcanic explosions. The rows are also ordered according to the duration of the climatic change; the top row refers to the shortest time interval.

Comparison between the observed and predicted values for the first two rows of Table 1 provides a test of the validity of our model calculations. The calculated increase in the stratospheric heating rate agrees well with the increase observed during the first few months after the Mt Agung eruption. Also, the calculated decrease in the average surface temperature for the 2-yr period following an eruption is in accord with the observed decrease to within their mutual uncertainties. The net effect of volcanic explosions is to cause a temporary cooling of the Earth.

Rows 3 and 4 of Table 1 permit an assessment of the contribution of volcanic activity to the ΔT observed over the past century. This comparison is shown in greater detail in Fig. 2, where we have also indicated the ΔT expected from the observed increase in the CO_2 content of the atmosphere, coming primarily from the burning of fossil fuels[22–24]. We have set the reference point of the temperature scale at 1885. The increase in T from the end of the nineteenth century until 1940, shown by the theoretical volcanic dust curve, arises from the declining level of volcanic activity over this period. Thus, the temperature was depressed during the early part of this period when there was much volcanic activity, but no such depression occurred during the later, volcanically quiescent times.

Figure 2 and Table 1 suggest that changes in the level of volcanic activity have made a significant contribution to the increase in surface temperature from the end of the nineteenth century to 1940, and a less significant contribution to the decline in temperature from 1940 to the present. Because we did not include possibly important feedback effects, such as variations in cloudiness, we cannot make a precise estimate

Table 1 Comparison between observed temperature changes and those expected from volcanic aerosols

Comparison		Characteristic of variable being compared			Model parameters				Comments
Observed value	Predicted value	Variable	Place	Time period	Aerosol species	$\Delta\tau$	F_1*	F_2†	
0.1 (ref. 34)	0.1‡	Atmospheric heating rate§	Southern Hemisphere, lower stratosphere	First several months after the Mt Agung eruption in 1963	silicate dust	0.1‖	—	—	The main source of the heating is absorption of outward-directed thermal radiation by the added aerosols rather than their absorption of solar radiation[4].
−0.3 (ref. 26)	−0.2	Mean surface temperature change, $\Delta T'$¶	Northern Hemisphere	Epochal average for a 2-yr period after an eruption	sulphuric acid	0 as against 0.1‖	0.25	1	The observed value was obtained by averaging the results of 13 separate events. This averaged value is statistically significant at the 1 % level.
+0.6 (ref. 35)	+0.5	Mean surface temperature change, ΔT¶	Northern Hemisphere	1880–1890 as against 1935–1945	sulphuric acid	0.07 as against 0'‖	1	1	
−0.3 (ref. 35)	−0.15	Mean surface temperature change, ΔT¶	Northern Hemisphere	1935–1945 as against 1965–1970	sulphuric acid	0 as against 0.02‖	1	1	The agreement between the predicted and observed ΔT from 1940 to 1970 is worse than is apparent from this Table. See Fig. 2.
∼+0.5 (refs 25, 33) +0.3–+0.6		Mean surface temperature change, ΔT¶	Northern Hemisphere	1450–1915 (Little Ice Age) as against 1935–1945	sulphuric acid	0.03** as against 0	1	1— 2	The full ice–snow albedo feedback will not be realised since there was insufficient time to build large continental ice sheet. However, observations suggest that some increase in snow cover may have occurred during the Little Ice Age[21].
+4 (refs 29, 33) +0.7–+7		Mean surface temperature change, ΔT¶	Both hemispheres	12,000– 25,000 BP as against now	sulphuric acid	0.02– 0.2†† as against 0	1	4	12,000–25,000 BP was the coldest part of the last ice age.

*Correction factor for the finite response time of the atmosphere–ocean–land system. A response time of 4 yr was used to calculate F_1 for the second row. This value was derived from the amount of mass in the mixed layer of the ocean[4,33].

†Correction factor for feedback from a change in ice and snow cover and in the atmospheric dynamics[17,18]. This correction factor is poorly known.

‡The tropospheric temperature was held constant because the response time of the troposphere is several months[4,36].

§In units of K d^{-1}.

‖Found from Fig. 1.

¶In K. ΔT as given in the first and second columns of the table is the difference in temperature between the second and first time period of the fifth column. For example, the observed value of +0.6 given in the third row means T increased by 0.6 K from 1880–90 to 1935–45.

**The $\Delta\tau$ value for the Little Ice Age was found from Lamb's[25] dust veil indices (DVI). The DVI for this period are based primarily on estimates of the amount of mass ejected by volcanic events. We assumed that DVI is proportional to $\Delta\tau$ and found the constant of proportionality from data on both quantities for the period 1880–1915 (ref. 4). The $\Delta\tau$ given in Table 1 for the Little Ice Age is thus only a rough estimate.

††The values of $\Delta\tau$ were derived from the sulphate concentration found in ice cores[28]. In deriving these values we assumed a residence time of 1 yr for the H_2SO_4 particles in the atmosphere, values of 30 cm yr^{-1} and 15 cm yr^{-1} for snow accumulation rates in Greenland and Antarctica, respectively[29,28], and the present 'normal' size distribution function of the stratospheric aerosols. The lower value of $\Delta\tau$ in the Table is a value found from the sulphate concentration for the Antarctic cores, while the higher value pertains to the sulphate concentration of the Greenland cores.

of these contributions. CO_2 was less important than volcanic aerosols in causing the temperature increase from the end of the nineteenth century until 1940. The relative ranking of these two effects, however, may well reverse itself during the next several decades and the Earth could experience a net warming over this period if no other factors intervene (such as man-made dust[16]).

During the period from ∼ 1450 to 1915, the Earth was enveloped in the 'Little Ice Age'. Mountain glaciers and sea ice were more extensive than now, Northern Hemisphere temperatures were ∼ 0.5 K colder than those for 1935–1945, and volcanic explosions were common[25–27]. The volcanic activity between 1880 and 1905 shown in Fig. 1 is typical of the Little Ice Age in terms of number and magnitude of eruptions[25]. The results of Table 1 suggest that the high level of volcanic activity during the Little Ice Age had an important role in causing the low temperatures of this period.

Ice cores from Greenland and Antarctica contain information on both aerosol content and temperature[28,29] that allow us to assess the importance of volcanic activity in generating and sustaining the last major ice age, the Wisconsin, which occurred

from ∼ 75,000 BP to 12,000 BP. Because the residence time of submicron-sized aerosols in the stratosphere is ∼ 1 yr, aerosols produced by a volcanic explosion are distributed uniformly over at least the hemisphere of the explosion, except for a short period immediately following the explosion[4,12]. Thus, aerosol-derived material found in ice cores provides a useful measure of the aerosol content of the stratosphere at the time the ice was layed down. Studies of the anion content of ice cores[28] show that their sulphate content was distinctly higher than its present value and reached a maximum value during the 13,000-yr period at the end of the Wisconsin, a time when T had its minimum value. The sulphate enhancement was significantly greater for the Greenland cores than for the Antarctic cores. Sulphate also occurs in tropospheric aerosols, so an increase could result from non-volcanic sources. Assuming that the enhancement in sulphate content is caused by volcanic aerosols, we obtain estimates of ΔT for the later part of the Wisconsin ice age (Table 1) and find that volcanic aerosols could have made a significant contribution to the observed low temperatures.

Table 2 Estimated effect of vehicles flying through the stratosphere

Aerosol source	Time frame	$\Delta\tau^*$,†	ΔT‡
SSTs and other high-flying aircraft	1990 nominal	4.9×10^{-4}	5×10^{-3}
	1990 maximum	8.9×10^{-4}	9×10^{-3}
	2000 nominal	1.1×10^{-4}	1×10^{-3}
	2000 maximum	3.3×10^{-4}	3×10^{-3}
Space Shuttles	1990	3.3×10^{-5}	1.5×10^{-4}

*The Space Shuttle and SST perturbations are estimates for the Northern Hemisphere. Southern Hemisphere values should be 3/7 as large. SSTs are responsible for 70% of the $\Delta\tau$ values.

†The $\Delta\tau$ values were obtained from estimates of the number of vehicles, their emission characteristics, and properties of the stratosphere[4,3,3].

‡Temperatures (K).

Further support for this proposition comes from the discovery[29] of ~ 2,000 volcanic dust bands and ~ 25 volcanic ash bands in Antarctic ice cores. The small particle sizes ($\leq 5 \mu m$) in the dust bands suggest that the dust originated from distant sources. The observed onset of sustained volcanic activity occurred simultaneously with observed strong cooling, the lowest temperature accompanied the greatest dust fall-out, and a cessation of dust bands and an increase in temperature were contemporaneous[29]. Finally, the average mass concentration of volcanic dust during the last part of the ice age implies the presence of a comparable amount of volcanic sulphate. This is consistent with the observed sulphate enhancement for the period.

While volcanic activity appears to have contributed to an intensification of the Wisconsin ice age during its later stages, the available evidence suggests that it had little to do with the early and middle portions of this ice age. Few dust bands occur in the Antarctic cores[29] for this period and little enhancement is seen in the sulphate content of the cores from both hemispheres[30].

At times earlier than the Wisconsin ice age, there are a few data that hint at a significant volcanic role in causing climatic change. A dramatic cooling known as the '90,000-yr BP' event has been detected both in polar cores and in sea cores. In the sea cores it was closely associated with a volcanic ash layer[30] and in the polar cores there was an order of magnitude more dust (not necessarily volcanic) during the event than at present[31]. Finally, studies of the number of volcanic ash layers in deep-sea
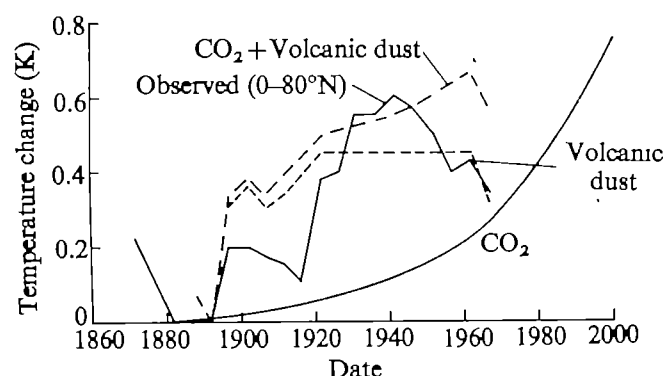
Fig. 2 Observed Northern Hemisphere surface temperature changes[25] (——) and theoretical estimates of the temperature changes from observed CO_2 and volcanic dust variations (·—·—·). CO_2 alone is the solid, steadily rising curve, and volcanic dust alone is – – – –. To estimate the volcanic temperature changes we use the optical depth data in Fig 1 averaged over 10-yr periods and only consider the years 1884–1892, 1903–1904, 1907–1908, 1912–1914 and 1963–1970 to have any volcanic dust present. The 10-yr averaging period was chosen to allow for response times in the Earth's climate system. An even better fit, especially between 1900 and 1920, could be obtained by other choices of averaging period.

cores indicate that the worldwide level of volcanic activity during the Quaternary, the time of the ice ages, was much greater than during the previous 20×10^6 yr (ref. 32).

In summary, volcanic aerosols seem to have made important contributions to climatic changes in the past, but apparently do not offer a complete explanation of all recorded changes.

Possible climatic impact of stratospheric flight

In a manner analogous to our treatment of volvanically produced stratospheric aerosols, we calculated the ΔT from increases in stratospheric particles from SSTs and Space Shuttles operating at projected traffic levels for the next several decades[6,7]. The results of these calculations are given in Table 2. Again, an increase in aerosol population in the stratosphere leads to a cooling of the Earth.

The importance of the ($\Delta\tau$, ΔT) values shown in Table 2 can be partially assessed by comparing them with $\Delta\tau$s produced by volcanic eruptions and ΔT values associated with significant climatic changes. The largest $\Delta\tau$ that SSTs may produce is an order of magnitude smaller than the $\Delta\tau$s produced by very minor volcanic activity, such as the Andes eruption of 1920 shown in Fig. 1. Thus, the contribution to the number of stratospheric aerosols by aircraft and rockets over the next few decades will be less than the contribution by random, quite insignificant volcanic eruptions.

According to Table 1, past climatic changes with $|\Delta T|$ values of 4 K and 0.5 K characterise the difference between glacial and interglacial periods and between the Little Ice Age and the decade 1935–1945, respectively. Clearly, a cooling of the Earth by 4 K would be disastrous for Mankind, since past ice ages have been characterised by huge sheets of ice covering almost all of Canada, most of the northern USA, and parts of northern and central Europe and the USSR. While it involved a much less dramatic change, the climate of the Little Ice Age was also significantly different from that of the present century. For example, the precipitation pattern in a number of marginal agricultural areas was notably different from the current one and the length of the growing season at high latitudes was significantly shorter[33]. Thus, past climatic changes of as little as 0.5 K have had a serious practical impact. Changes of < 0.1 K would be less serious, however, than those that have occurred over the past several decades and would even be hard to measure. All the computed ΔT values of Table 2 are below this bound of 0.1 K; hence we conclude tentatively that the aerosols produced by vehicular traffic in the stratosphere over the next several decades will not have serious climatic consequences.

This conclusion can be put into some perspective by noting that about 70 Space Shuttle flights per day (against projections of < 1 per week) or 5,000 SSTs in operation (against projections of 250–1,500) would be required to decrease the surface temperature by 0 1 K. The conclusion thus reached on Space Shuttles is unlikely to be altered by future data, but the conclusion on the effect of high-flying aircraft should be re-evaluated as projections concerning their emission and flight characteristics are updated.

Finally, we note that the presence of additional particles in the stratosphere will have a negligible effect on the ozone cycle. For example, we estimate that the fractional increase in the ozone destruction rate due to scattering by the added particles is only ~ 0.1 of the $\Delta\tau$ values shown in Table 2 (ref. 7).

1 *CIAP monograph 1*, Department of Transportation document DOT-TST-75-51 (1975).
2 Lazrus, A., Gandrud, B., and Cadle, R. D., *J. geophys. Res.*, **76**, 8083 (1971).
3 Toon, O. B., and Pollack, J. B., *J. geophys. Res.*, **78**, 7051 (1973).
4 Pollack, J. B., et al., *J. geophys. Res.*, **81**, 1071 (1976).
5 Mossop, S. C., *Nature*, **203**, 824 (1964).
6 *CIAP Report of Findings*, Department of Transportation document DOT-TST 75-50 (1974).

7 Pollack, J. B., Toon, O. B., Summers, A., Van Camp, W., and Baldwin, N., J. appl. Meteorol., 15, 247 (1976).
8 Pollack, J. B., and Toon, O. B., U.S. Dept. Transportation document DOT-TSC-OST-74-15, 457 (1974).
9 Lacis, A. A., and Hansen, J. E., J. atmos. Sci., 31, 118 (1974).
10 Pollack, J. B., Icarus, 10, 301 (1969).
11 Toon, O. B., and Pollack, J. B., J. appl. Meteorol., 15, 225 (1976).
12 Atwater, M. A., Science, 170, 64 (1970).
13 Pollack, J. B., Toon, O. B., and Khare, B. N., Icarus, 19, 372 (1973).
14 Palmer, K., and Williams, D., Appl. Optics, 14, 208 (1975).
15 Toon, O. B., Pollack, J. B., and Khare, B. N., J. geophys. Res. (in the press).
16 Rasool, S. I., and Schneider, S. H., Science, 173, 138 (1971).
17 Budyko, M. I., Soviet Geogr., 10, 429 (1969).
18 Sellers, W. D., J. appl. Meteorol., 12, 241 (1973).
19 Barry, R. G., Ives, J. D., and Andrews, J. T., Quat. Res., 1, 415 (1971).
20 Lamb, H. H., and Woodroffe, A., Quat. Res., 1, 29 (1970).
21 Barry, R. G., Andrews, J. T., and Mahaffy, M. A., Science, 190, 979 (1975).
22 Manabe, S., and Wetherald, R. T., J. atmos. Sci., 24, 241 (1967).

23 Manabe, S., and Wetherald, R. T., J. atmos. Sci., 32, 3 (1975).
24 Machta, L., Bull. Am. met. Soc., 53, 402 (1972).
25 Lamb, H. H., Phil. Trans. R. Soc., A266, 425 (1970).
26 Bray, J. R., Adv. ecol. Res., 7, 177 (1971).
27 Lamb, H. H., Climate Present, Past and Future (Methuen, London, 1972).
28 Cragin, J. H., Herron, M. M., Langway, C. C., and Klouds, G., SCOR/SCAR Polar Oceans Conf., McGill University, Montreal, Canada (1974).
29 Gow, A. J., and Williamson, T., Earth planet. Sci. Lett., 13, 210 (1971).
30 Kennett, J. T., and Huddleston, P., Quat. Res., 2, 384 (1972).
31 Dansgaard, W., and Hammer, C. U., Geol. Soc. Am., 6, 1014 (1973).
32 Kennett, J. P., and Thunell, R. C., Science, 187, 479 (1975).
33 Bryson, R. A., Science, 184, 753 (1974).
34 Newell, R. E., J. atmos. Sci., 27, 977 (1970).
35 Mitchell, J. M., in Global Effects of Environmental Pollution (edit. by Singer, F.), 139 (Springer Verlag, New York, 1970).
36 Gierasch, P., Goody, R., and Stone, P., Geophys. Fluid Dyn., 1, 1 (1970).
37 Castleman, A. W., Jr, Munkelwitz, H. R., and Manowitz, Tellus, 26, 222 (1974).

The toxicity of ^{90}Sr, ^{226}Ra and ^{239}Pu

M. C. Thorne & J. Vennart

MRC Radiobiology Unit, Harwell, Didcot, Oxfordshire OX11 0RD, UK

Data now available on the risks of radiation-induced fatal cancer and hereditary disease and radionuclide metabolism suggest that limits on the rates of intake of ^{90}Sr, ^{226}Ra and ^{239}Pu at work, presently recommended by the International Commission on Radiological Protection, might be in need of considerable revision one with another and with the limit for uniform exposure of the whole body.

DISCUSSIONS on this subject often centre on the values of maximum permissible body burden (MPBB) recommended by the International Commission on Radiological Protection (ICRP)[1]. For example, Spiers and Vaughan[2] considered the risk of osteosarcoma, cancer of the air sinuses and leukaemia arising from ^{226}Ra and ^{239}Pu in the skeleton, and they concluded that the currently recommended values of MPBB for these radionuclides do not need major revision. As Spiers and Vaughan[2] acknowledged, however, the recommendations of ICRP limit the rate of intake of radioactive materials into the body so that the MPBB for a well retained, long lived radionuclide will not be reached until the end of a working lifetime of 50 yr. In such circumstances, MPBB might be regarded more as an index by which to judge whether an individual has been overexposed, than as a quantity useful in planning for work with the radionuclide concerned. For this latter purpose, the corresponding maximum permissible annual intake (MPAI) is more pertinent. Values of MPAI may be obtained from the values of maximum permissible concentration in air or water (MPC) recommended by the Commission[1]. It is worth noting that, even if there were no debate about the relationship between risk of an effect and dose, there might now be a need to revise values of MPAI because of the data on the uptake and retention of radionuclides in body tissues which have become available since the recommendations were made.

The recommended values of MPAI for each radionuclide are different for soluble and insoluble compounds and for inhalation and ingestion. Using methods similar to those described in the MRC report on the toxicity of plutonium[3], we have estimated by way of example, the total risk of fatal cancer and serious hereditary disease arising from currently recommended values of MPAI for ingested soluble ^{90}Sr and ^{226}Ra and inhaled insoluble ^{239}Pu. We show that in these three cases the values of MPAI may be in need of considerable revision.

Relationships between risk and dose

Data on the risk of radiation-induced fatal cancer and hereditary disease have been reviewed by a United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR)[4] and a National Academy of Sciences Advisory Committee on the Biological Effects of Ionizing Radiations (BEIR)[5]. Both reports adopt the hypothesis that the risk of an effect is linearly related to dose. We have used the information in these reports to obtain the values discussed below and because they are based on a hypothesis which cannot yet be proved at the low doses concerned in radiological protection, they must be regarded as 'nominal' risks.

Leukaemia. Epidemiological studies on the survivors of the A-bomb explosions at Hiroshima and Nagasaki show an increasing leukaemia incidence with increasing dose. From the BEIR report[5], the estimated incidence is 20–50 cases per 10^6 people exposed to 1 rad ($20-50/10^6$/rad), in good agreement with the UNSCEAR estimate[4] of 15–40 cases/10^6/rad and the estimate of Smith and Doll[6] of about $20/10^6$/rad. We have chosen a value of $30/10^6$/rad for radiation of low linear energy transfer (LET).

Breast cancer. Data on breast cancer incidence in women survivors of the A-bombs and in tuberculosis patients subjected to multiple X-ray fluoroscopies show an incidence rate of between 3 cases/10^6/yr/rad (UNSCEAR) and 6 cases/10^6/yr/rad (BEIR). We have assumed a 20-yr period of increased incidence to yield a morbidity of about 90 cases/10^6/rad. Since the 5-yr survival of patients treated for breast cancer is about 50% (ref. 4) we have chosen a rounded value for the incidence of fatal breast cancer in females of $50/10^6$/rad of low LET radiation.

Lung cancer. The UNSCEAR report gives an estimate of 10–40 cases/10^6/rad of high dose rate γ radiation from A-bombs during the first 25 yr after exposure. This is in good agreement with the value in the BEIR report of 0.9 cases/10^6/yr/rem obtained from groups exposed to A-bomb radiations and to X rays in the treatment of ankylosing spondylitis. The value given in the BEIR report for exposure of the respiratory system to α rays from radon is 11.2 cases/10^6/yr/rad for a follow-up period of about 25 yr. Using a value of 20 for the quality factor for α rays (*vide infra*) this corresponds to a total risk of about 14 cases/10^6/rad of low LET radiation. Based on these two sets of data we have used a value of $20/10^6$/rad for low LET radiation.

Thyroid cancer. The UNSCEAR report estimates the incidence rate for thyroid cancer in adults to be in the range of 1 to 4 cases/10^6/yr/rad and the BEIR report does not give a numerical estimate. We have assumed a value of 2.5 cases/10^6/yr/rad and a 20-yr period of increased incidence, to give a total morbidity of 50 cases/10^6/rad. The 5-yr survival from thyroid cancer, however, is about 80% (ref. 4) and we have chosen a value $10/10^6$/rad of low LET radiations for the risk of radiation-induced fatal thyroid cancer.

Osteosarcoma. There is little information on the induction of osteosarcoma by external radiation. From data on patients treated with X rays for ankylosing spondylitis, the BEIR report gives an estimate of 0.034–0.22 cases/10^6/yr/rad average dose to bone over a follow-up period of 5–27 yr. The BEIR report also gives an estimate of 0.11 cases/10^6/yr/rem average dose to bone for persons contaminated with ^{226}Ra. This value is based on a

quality factor of 10 for α particles and a follow-up period of 49 yr. It is, therefore, approximately equivalent to a risk of 110 cases/10⁶/α rad to cells on the endosteal surface of bone which receive about half the average dose in bone. This is in good agreement with the estimate by Spiers and Vaughan[2] of 95 cases/10⁶/α rad to the endosteum. We have, therefore, taken a risk estimate of 100/10⁶/rad to the endosteum for high LET radiation and 5/10⁶/rad for low LET radiation, the latter being consistent with data from the BEIR report and the quality factor of 20 for α particles (*vide infra*).

Air-sinus carcinoma. In persons contaminated with ²²⁶Ra the risk of air-sinus carcinoma induction is about 40% of that of osteosarcoma[2], however, no such cancers have been induced in patients treated with ²²⁴Ra (ref. 5). These data suggest that the air-sinus carcinomas arising in persons contaminated with ²²⁶Ra are a result of irradiation of the air-sinus epithelium by ²²²Rn and its short lived radioactive daughters trapped in the air cavities of the sinuses[7]. Therefore, we have assumed the risk of air-sinus cancer to be zero in persons exposed to ⁹⁰Sr and ²³⁹Pu and to be 40% of the risk of osteosarcoma in persons exposed to ²²⁶Ra (compare Table 3).

Other cancers. In the A-bomb survivors, an excess of cancers other than those listed here has been observed. This excess is not sufficient, however, to break down into specific categories. The estimate for these 'other' cancers by UNSCEAR was about 40 cases/10⁶/rad during the 25 yr after exposure and by BEIR about 2.0/10⁶/yr/rem. We have used a value 40/10⁶/rad of low LET radiation for the combined total of all 'other' fatal cancers. In the case of inhaled insoluble ²³⁹Pu, the only 'other' tissue irradiated to any significant extent is the liver. For this organ we have used a risk estimate of 10/10⁶/rad of low LET radiation, derived from the MRC report[3] when allowance is made for the different values of the quality factor used here and by MRC.

Serious hereditary disease. The risk of genetic effects from irradiation of the gonads is particularly difficult to estimate. Both the BEIR and UNSCEAR reports considered that the risk from irradiating oocytes is considerably less than that from irradiating spermatagonia. UNSCEAR derives a best value of 20 cases of serious hereditary disease per 10⁶ in the immediate descendants of a male parent exposed to 1 rad of low LET radiation at low dose rate. Comparable estimates in the BEIR report for 1 rem to both parents, were 10–100 cases/10⁶ in the first generation and 50–500 at equilibrium. BEIR also estimated the increase in congenital anomalies, anomalies expressed later and constitutional and degenerative diseases after both parents had received 1 rem, to be 1–100 cases/10⁶ in the first generation and 10–1,000/10⁶ at equilibrium. For workers exposed to radiation, the relevant risk is probably that of serious hereditary disease in their immediate descendants and for this we have chosen a value of 50 cases/10⁶ for a worker receiving 1 rad of low LET radiation. The appropriate value for irradiation of the general population, if evaluated over many generations, would be greater, and greater still if diseases of less severity were taken into account.

Table 1 'Nominal' risk of fatal cancer and serious hereditary disease per unit dose

Effect	Tissue at risk	Cases/10⁶/rem Male workers	Cases/10⁶/rem Female workers
Leukaemia	Red bone marrow	30	30
Breast cancer	Female breast	—	50
Lung cancer	Lung	20	20
Thyroid cancer	Thyroid	10	10
Osteosarcoma	Cells on endosteal surfaces	5	5
Liver cancer	Liver	10	10
Other cancers	All other tissues	30	30
Hereditary disease	Gonads	50	50
Total	Whole body	155	205

The various risk estimates discussed above are summarised in Table 1 and are given per rem. In the calculations of dose equivalent discussed below it will be assumed that a quality factor (*Q*) of unity applies for low LET radiations and 20 for α particles as recommended[8]. The value of 20 for α particles is supported by cytogenetic work with ²³⁹Pu (ref. 9) and experiments on chromosome damage in liver cells with low dose rate radiation[10]. A value of 20 is also compatible with the data on osteosarcoma induction mentioned above. ICRP used a value of 10 in its 1959 recommendations[1], as did the MRC[3].

Dose commitment per unit activity taken into the body

Since it is assumed that risk of an effect is linearly related to dose, risk in any organ or tissue from intake of radionuclide is proportional to the resulting average total dose received by that tissue (the dose commitment) before induction of the effect. The tissues considered to be at risk are shown in Table 1. Their masses have been taken from the data given by ICRP for Reference Man[11]. Metabolic models for the three radionuclides under discussion have been proposed in the reports of two ICRP Task Groups[12,13].

For inhaled insoluble ²³⁹Pu we have used the lung model proposed in ref. 12 to calculate the dose commitment in lung (comprising the trachea, bronchi, pulmonary region and associated lymph nodes) and the fractions of ²³⁹Pu transferred to the gut and the blood. We have assumed the inhaled particles to have an activity median aerodynamic diameter (AMAD) of 1 μm, however, differences in particle size within the range likely to be encountered in practice will not markedly affect the results[3]. The dose to the walls of the gut is small compared with that in other tissues because of the limited range of the α particles and the position of sensitive cells[3]. Of ²³⁹Pu entering the blood, a fraction (0.45), is assumed to reach both bone and liver and be retained with half lives of 100 and 40 yr respectively[12]. From data of Richmond and Thomas[14] we have estimated that of ²³⁹Pu entering blood 10⁻³% is deposited per g of testes or ovaries and remains there indefinitely[15], and from experimental work in our unit, the dose to the spermatogonial stem cells is taken to be 2.5 times the average dose in the testes[15]. Our experimental work on female mice so far has not suggested that the dose to maturing oocytes is very different from the average in the ovary.

Average doses to cells on endosteal surfaces (taken to lie within 10 μm of the bone surface) and to haematopoietic bone marrow have been calculated for ²³⁹Pu and ²²⁶Ra using the Monte Carlo techniques described by Thorne[16]. ²³⁹Pu was assumed to be distributed uniformly in a thin layer over all endosteal surfaces and ²²⁶Ra in the volume of mineral bone (details of these calculations will be published elsewhere). Doses to the endosteum and red bone marrow from ⁹⁰Sr have been calculated using the data of Whitwell and Spiers[17].

For ingested soluble ⁹⁰Sr and ²²⁶Ra, the fractional absorption from the gut has been taken as 0.3 and 0.2 respectively[13]. The value for ⁹⁰Sr is a compromise between the value 0.2 given in ref. 13 and the mean value 0.38 given in ref. 11. Values for the time integral of activity of ⁹⁰Sr and ²²⁶Ra in trabecular and compact bone and in the soft tissue of the body have been derived using data in ref. 13. ⁹⁰Sr is assumed to be in equilibrium with its short lived daughter ⁹⁰Y in all organs and tissues of the body at all times after ingestion. All ²²²Rn produced from ²²⁶Ra in the soft tissues of the body is assumed to escape from the body before it decays. Of ²²²Rn produced in bone, 30% is assumed to remain at its production site, where it gives up the α energy of all its daughters except that of ²¹⁰Po, which follows the long lived ²¹⁰Pb. The remaining 70% is assumed to escape from the body without imparting any energy to tissues[1]. For both ingested soluble ⁹⁰Sr and ²²⁶Ra, the dose to the gut is very much smaller than that to the other organs and tissues shown in Table 2 and it has been neglected.

Values are given in Table 2 for the dose commitments in various organs and tissues of the body from ingestion of 1 μCi

Table 2 Dose commitments in various organs and tissues of the body from ingestion of 1 µCi of soluble [226]Ra, ^{90}Sr and inhalation of 1 µCi of insoluble [239]Pu

	Dose commitment (rem µCi⁻¹)		
Tissue at risk	Ingested [226]Ra	Ingested [90]Sr	Inhaled [239]Pu
Red bone marrow	2	0.7	300
Female breast	0 4*	0.005*	—
Lung	0.4	0.005	1200
Thyroid	0.4	0.005	—
Cells on endosteal surfaces	20	0.9	3700
Liver	0.4*	0.005*	760
All other tissues	0.4	0.005	—
Gonads: Male	0.4	0.005	38
Female	0.4	0.005	15

* [90]Sr and [226]Ra are assumed to be uniformly distributed in soft tissues[11,12].

of soluble [226]Ra and ^{90}Sr and from inhalation of 1 µCi of insoluble [239]Pu. For the gonads, dose has been integrated over an assumed reproductive life of 20 yr and for other tissues, over a working life of 50 yr. These dose commitments therefore only apply to intakes which occur early in the working life.

Risk per unit activity taken into the body

Values of the 'nominal' risk per unit intake of the radionuclides were derived from the data in Tables 1 and 2 and are shown in Table 3. The risk estimates apply to intakes early in life and are overestimates of the true risk, in the sense that some cancers induced late in life will not be recognised because of their latent periods. Several aspects of the risks shown in Table 3 are worth noting.

(1) For [226]Ra a risk of leukaemia, about half that of osteosarcoma, seems at variance with human experience: in people containing the radionuclide, the incidence of leukaemia has been about ten times' lower than that of osteosarcoma[2]. A plausible explanation for this disparity might be that the use of an arbitarily defined quality factor of 20 for α particles[8] overestimates the risk of leukaemia relative to that of osteosarcoma. An alternative explanation concerns the distribution of cells sensitive to leukaemia induction within red bone marrow and the relationship between risk of an effect and dose for those cells. Cells in marrow at distances greater than the range of α particles (about 40 µm) from bone receive no dose from α particles arising in bone and cells nearer to bone receive increasing doses the nearer they are. Our estimates of the risk of leukaemia have been obtained from the estimated average dose in marrow and the observed risk of leukaemia per unit dose to the marrow of people who were uniformly irradiated either with A-bomb radiations or X rays. Therefore, our estimates will be in error if the sensitive cells are not uniformly

distributed in the marrow or if the dose–response relationship is not linear over the range of doses concerned. With regard to the latter, the lifetime doses to marrow which are of interest in radiological protection are within the range from which risk estimates were obtained[4,5]. It is possible, however, that the assumption of linearity overstates the risk at low dose rates. Conversely, if a significant proportion of sensitive cells lie near to bone, we might have underestimated the risk of leukaemia because such cells receive doses much greater than the average in the marrow. We have experimental evidence that the incidence of myeloid leukaemia in mice reaches a maximum value at average doses in marrow of about 200–300 rad of X rays (I. Major and R. H. Mole, personal communication) presumably because at greater doses, the sterilisation of transformed cells exceeds their production. Smith and Doll[5] have also drawn attention to this breakdown of the linear dose–response hypothesis to explain the absence of leukaemia in women treated for cancer of the cervix who received very large doses to small volumes of marrow[18]. In the radium cases exhibiting osteosarcoma[2], marrow cells near to bone received very large doses. If a significant proportion of marrow cells sensitive to leukaemia induction are near to bone, this might explain the low incidence of leukaemia observed. By contrast, it is interesting that osteosarcoma incidence in these radium cases was increasing at α-particle doses to cells within 10 µm of endosteal surfaces of 500 rad (10,000 rem) or more[19]. Clearly, more information is required about the distribution of the cells which are sensitive to induction of both osteosarcoma and leukaemia.

(2) For [226]Ra the predicted risk of hereditary disease is about 8% of the total risk, corresponding values for ^{90}Sr and [239]Pu are 1 and 2% respectively. These percentages would increase for populations exposed to the radionuclides for long periods.

(3) The risk of leukaemia predominates for [90]Sr. The relative risk of leukaemia to osteosarcoma for both ^{90}Sr and [239]Pu is much greater than would be expected from experiments on animals. Most of these experiments have, however, been conducted in ways which might not have favoured the incidence of myeloid leukaemia as discussed in (1) above.

(4) Although all three radionuclides are commonly referred to as "bone-seekers", these predictions show that risk arises mainly from irradiation of the skeleton and its marrow only in the case of ingested [90]Sr (98% of the total); for ingested [226]Ra, the risk from irradiating the skeleton is about 65% and for inhaled [239]Pu, about 45% of the total. In the last case, about 40% of the risk arises from irradiation of the lung.

We acknowledge that the estimates of 'nominal' risk we have made are open to criticism and that the data on which they are based have been interpreted in different ways (see, for example, ref. 20). There is also considerable uncertainty concerning estimates of dose and the relevant tissues at risk. Nevertheless, the estimates of 'nominal' risk (Table 3) were derived from the most recent data in official reports and it is interesting to

Table 3 'Nominal' risk per unit activity of soluble [226]Ra and ^{90}Sr ingested and insoluble [239]Pu inhaled, for male and female workers

	Cases/10⁶/µCi					
Effect	Ingested [226]Ra		Ingested [90]Sr		Inhaled [239]Pu	
	MW	FW	MW	FW	MW	FW
Leukaemia	60	60	21.00	21.00	9,000	9,000
Breast cancer	—	20	—	0.25	—	—
Lung cancer	8	8	0.10	0.10	24,000	24,000
Thyroid cancer	4	4	0.05	0.05	—	—
Osteosarcoma	140*	140*	4.5	4.5	18,500	18,500
Liver	4	4	0.05	0.05	7,600	7,600
Other cancers	12	12	0.15	0.15	—	—
Hereditary disease	20	20	0.25	0.25	1,900	750
Total	248	268	26.1	26.35	61,000	59,850

* Increased from 100 to 140 to allow for the risk of air-sinus carcinoma in persons exposed to [226]Ra.
MW, Male workers; FW, female workers.

Table 4 Annual limits on exposure recommended by ICRP; the resulting values of MABB after exposure for 50 yr and the risk from 1 yr of practice exposed to the limits

Exposure	Annual limit	MABB (μCi)	'Nominal' risk (cases per 10,000 per year)
Ingested soluble ^{226}Ra	0.08 μCi	0.005	0.2
Ingested soluble ^{90}Sr	0.8 μCi	0.37	0.2
Inhaled insoluble ^{239}Pu	0.07 μCi	0.16	40
Irradiation of whole body	5 rem		9

consider the consequences of applying these estimates to current recommendations of the ICRP.

Body burdens and risk from the MPAIs recommended by the ICRP

The values of MPAI recommended at present by the ICRP for the radionuclides under discussion are shown in Table 4. They were obtained directly from the values of MPC recommended by ICRP, assuming that a Reference Man[11] takes in 2,200 cm^3 of water and inhales 2×10^7 cm^3 of air each day. Table 4 shows the maximum achievable body burden (MABB) which such a man could attain at the end of 50 yr if he were exposed to the MPAI each year for that period. The values for ^{90}Sr and ^{226}Ra of 0.37 and 0.005 μCi respectively have been obtained using the data given in ref. 13. About 90 and 80% of the ^{90}Sr and ^{226}Ra respectively, would be in bone and the remainder in soft tissues. These values of MABB are about 5 and 20 times less than those of MPBB recommended at present[1], namely 2 and 0.1 μCi for ^{90}Sr and ^{226}Ra respectively, of which 99% was assumed to be in bone. Values of MPC, and therefore of MPAI, were calculated to result in values of MPBB after a 50-yr exposure and the ratio of the presently estimated values of MABB to the recommended values of MPBB demonstrates the effect of using the new metabolic data now available[13]. No value of MPBB was given by ICRP[1] for inhaled insoluble ^{239}Pu, its intake being limited by the dose in lung. The value of MABB shown in Table 4 (0.16 μCi) was derived using the data from ref. 12. About 40% of this MABB will be in bone, about 30% each in liver and lung and only a small percentage in other soft tissues, including the gonads. The value of MPBB commonly used for ^{239}Pu (0.04 μCi) was derived by assuming that 90% of the activity would be in bone and is four times smaller than the value of MABB calculated here.

The average 'nominal' risks for men and women arising from the MPAIs are also shown in Table 4 and, for comparison, that arising from the dose limit of 5 rem to the whole body[20] recommended at present, assuming a 'nominal' risk of 180 cases per 10^6 per rem, which is a representative value for men and women (Table 1). Since it is reasonable to expect that the risk from exposure to each of the various limits should be the same, the three values of MPAI and the dose limit for whole-body exposure are clearly in need of revision one with another. If it were agreed that the risk from exposure of the whole body

to 5 rem each year is acceptable to workers, then the MPAI for inhaled insoluble ^{239}Pu would have to be revised down by a factor of about five and those of ^{90}Sr and ^{226}Ra both revised up by a factor of about 50. The ICRP is at present revising its recommendations and will no doubt ensure that the hazards from occupational exposure will not exceed those that are accepted in most other industrial or scientific occupations with a high standard of safety (para. 47 of ref. 21). In this respect, the MRC report[3], quoting a paper by Pochin[22], compared the risk from an MPAI with the average risk of fatalities in other kinds of work for which there seems to be a dividing line at 1 in 10,000 per year between the relatively safe and less safe industries. Comparisons of this kind need to take account of many social and economic factors, discussion of which is outside the scope of this paper. Whatever may be the outcome of discussions on these matters, however, it seems likely that many values of MPAI will need to be revised, not always in the same direction, because of new data on the fate of radioactive materials after they have entered the body.

1 *Report of Committee 2 on Permissible Doses for Internal Radiation* (Pergamon, Oxford, 1959).
2 Spiers, F. W., and Vaughan, J., *Nature*, **259**, 531 (1976).
3 *The Toxicity of Plutonium* (HMSO, London, 1975).
4 UNSCEAR, *Ionising Radiation: Levels and Effects* (United Nations, New York, 1972).
5 BEIR, *The Effects on Populations of Exposure to Low Levels of Ionising Radiation* (National Academy of Sciences, National Research Council, Washington, 1972).
6 Smith, P. G., and Doll, R., *Br. J. Radiol.*, **49**, 224 (1976).
7 Aub, J. C., Evans, R. D., Hempelmann, L. H., and Martland, H. S., *Medicine*, **31**, 221 (1952).
8 *Data for Protection against Ionising Radiation from External Sources: Supplement to ICRP Publication 15* (Pergamon, Oxford, 1973).
9 Beechey, C. V., Green, D., Humphreys, E. R., and Searle, A. G., *Nature*, **256**, 577 (1975).
10 Brooks, A. L., *Science*, **190**, 1090 (1975).
11 *Task Group Report on Reference Man* (Pergamon, Oxford, 1975).
12 *The Metabolism of Compounds of Plutonium and the Other Actinides* (Pergamon, Oxford, 1972).
13 *Alkaline Earth Metabolism in Adult Man* (Pergamon, Oxford, 1973).
14 Richmond, C. R., and Thomas, R. L., *Hlth Phys.*, **29**, 241 (1975).
15 Green, D., Howells, G. R., Humphreys, E. R., and Vennart, J., *Nature*, **255**, 77 (1975).
16 Thorne, M. C., *Nature*, **259**, 539 (1975).
17 Whitwell, J. R., and Spiers, F. W., *Phys. Med. Biol.*, **21**, 16 (1976).
18 Hutchinson, G. B., *J. natn. Cancer Inst.*, **40**, 951 (1968).
19 Rowland, R. E., Keane, A. T., and Lucas, H. F., Jr, in *Radionuclide Carcinogenesis, AEC Symp. Ser. No. 29* (edit. by Sanders, C. L., *et al.*), 406–420 (AEC, Washington, 1973).
20 Mole, R. H., *Br. Inst. Radiol.*, **48**, 157 (1975).
21 *Recommendations of the International Commission on Radiological Protection, Adopted September 1965* (Pergamon, Oxford, 1966).
22 Pochin, E. E., *Community Health*, **6**, 2 (1974).

Darwin's finches and the evolution of sexual dimorphism in body size

Jerry F. Downhower

Department of Zoology, The Ohio State University, Columbus, Ohio 43210

A simple model based on certain bioenergetic phenomena accounts for seasonal patterns of sexual dimorphism, size related changes in the timing of reproduction and geographic variation in body size of Darwin's finches.

SEXUAL dimorphism in size influences the biology of birds in many ways; for example, differences in beak size determine the size and kind of food that a bird can capture and ingest[1-3]. Consequently, monogamous pairs that differ in size may increase the rate at which they feed their nestlings, since the adults do not compete for the same food items[1,2]. Interspecific aggression[4]

or intraspecific aggressive conflicts between males, where increased size could influence the outcome, may account for the general avian condition that males are the larger sex[2]. In addition, sexual differences in size may facilitate pair formation[5,6].

Many observations indicate that sexual dimorphism in body size may be subject to another equally broad explanation. First, among monomorphic species, sexual differences in behaviour result in each sex foraging on different things in different places[7-10]. Second, full adult size is generally achieved at fledging, and certainly before first reproduction, so increased male size in polygynous species may reduce the risks associated with delayed sexual maturation in males[11]. Third, sexually dimorphic species in which the female is the larger sex require special explanations[12-14]. Finally, since beak and body size are correlated, dietary differences associated with beak size may be an inevitable consequence of the advantages associated with size alone.

Here I examine the proposition that sexual dimorphism in body size influences the timing of reproduction in fluctuating and seasonal environments. Evidence favouring the hypothesis is drawn from general studies on the energetic consequences of differences in body size, Lack's studies on morphological variation in Darwin's finches, and from my own studies of the reproductive biology of *Geospiza conirostris* and *G. fuliginosa*, two seed-eating finches that inhabit Española, the southernmost of the Galápagos Islands.

The influence of body size

Body size (weight) influences the rate at which birds use energy. Resting metabolism, the energetic costs of flight[16], and existence metabolism[17] bear an exponential relationship to body weight as do egg weight[18], heart rate and conductance[19]. In general, the exponent of these relationships is less than 1.0. In contrast, functions describing the storage capacity or energy reserves of a bird, for example, digestive capacity, and lipid reserves[20],-bear a nearly linear relationship to body weight. These different relationships affect the ability of a bird to respond to periods of deprivation and surfeit.

Consider two birds differing only in weight that are subjected to a physiological stress such as fasting or water deprivation after having been maintained on an *ad libitum* diet. To withstand this period of stress, they draw on stored reserves. The rate at which the larger (R_L) or the smaller (R_s) individual uses energy is dependent on body weight and can be expressed as follows

$$R_L = aW_L{}^b$$
and
$$R_s = aW_s{}^b$$

where W_L and W_s are the weights of the larger and smaller individuals. Since a and b are constants, R_L is always greater than R_s.

How long (T_L, T_s) each individual persists, depends on the size of its reserve ($cW_L{}^d$ and $cW_s{}^d$) and the rate at which energy is used

$$T_L = cW_L{}^d/R_L = \frac{c}{a} W_L{}^{(d-b)}$$

and

$$T_s = cW_s{}^d/R_s = \frac{c}{a} W_s{}^{(d-b)}$$

where c and d are constants. As long as d is greater than b, T_L will be greater than T_s and the larger individual will persist longer than the smaller. It also follows, that when confronted with a surfeit of resources, the smaller individual can replenish its reserves faster than the larger ($R_L > R_s$), provided that size does not influence the efficiency of uptake in any disproportionate way.

This analysis of the consequences of differences in body size is inherent in explanations of geographical variation in size of Atlantic salmon (*Salmo salar*)[20], cricket frogs (*Acris crepitans*)[21],
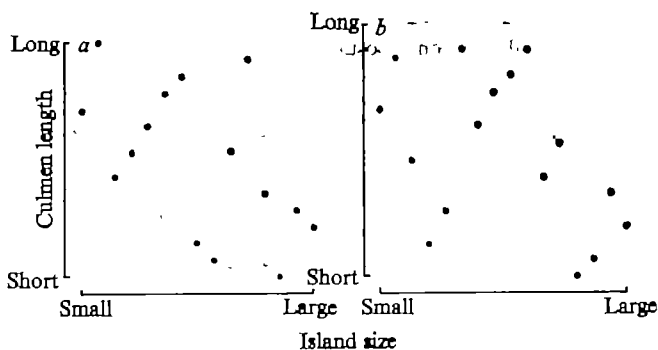
Fig 1 Rankings of culmen length and island size for populations of (a) *Geospiza fuliginosa* and (b) *Certhidea olivacea*. Culmen length data are taken from Lack[15]; island sizes are given by Bowman[28]. a, $R_s = -0.55$, $P < 0.05$, $n = 15$; b, $R_s = -0.44$, $P < 0.05$, $n = 16$.

wood rats (*Neotoma* sp.)[22], and has been expanded in detail for birds by Calder[19].

In fluctuating environments, consequences of differences in body size may become important. Among temperate species, periods of inclement weather alter foraging behaviour[23], and severe weather conditions can alter the survivorship of conspecifics that differ in body size[24,25].

Darwin's finches and the Galápagos

The lowlands of the Galápagos Islands are dry, hostile places, but on the larger islands these 'arid zone' forests give way to more mesic environments at higher elevations[26] and on these islands, Darwin's finches migrate from the arid coast to the highlands during periods of drought[26,27]. On smaller, low lying islands, migration may be useless since the entire island may be covered with arid zone vegetation, so the finches are more likely to be subjected to periodic stress than would conspecific populations on larger islands. It follows from the analyses of the consequences of body size that finches on small islands should be larger than conspecifics on large islands.

Geospiza fuliginosa, the smallest of the seed-eating Geospizinae, occurs on many islands in the Galápagos, as does *Certhidea olivacea* the smallest insectivorous finch in the archipelago[17,18]. In both cases, the birds on small arid islands tend to be larger than their conspecifics on larger islands (Fig. 1). Although this trend is consistent with the conclusion that increased size provides some resistance to environmental change, it is also consistent with the idea that size is determined by diet[2], since both seeds[29] and insects[30] are likely to be larger in more arid environments. The size and frequency distributions of seeds and insects on islands in the Galápagos are not known, but the effects of differences in prey size and climatic variability are complementary, in that both would select for larger size on smaller islands.

Body size and timing of reproduction

The reproductive patterns and life histories of vertebrates are also influenced by differences in body size. Large species generally mature slowly, produce small clutches, and are long lived, whereas smaller species mature more rapidly, produce larger clutches and are shorter lived[31]. The possibility that differences in body size would influence the timing of reproduction in seasonal environments has not been investigated.

Consider two species, with similar ecologies and competitive abilities, that differ in body size. On the assumptions that reproduction is initiated by a variable environmental cue and that the energy for egg production is derived from currently available food supplies rather than from stored reserves, it follows that females of the smaller species are likely to accumulate the necessary energy for egg production sooner than those of the larger species because they have a reduced maintenance cost[19]. The smaller species should, therefore, breed earlier than the larger species (Fig. 2a).

Fig 2 A graphic model for size-dependent changes in the timing of reproduction. Sm, Smaller species; Lg, larger species. See text for discussion.

its breeding population is at a maximum level (Fig. 2b) and shifts into a nestling phase (nestlings present in more than half the nests) before the population of the smaller species does so (C in Fig. 2c). Since the smaller species can establish nests until somewhat later, it will not achieve maximum densities until then (Fig. 2b), and so it enters the nestling phase (D in Fig. 2c) later than the larger species.

My studies of the reproductive biology of two species of Darwin's finches, *Geospiza fuliginosa* and *G. conirostris*, provide a direct test of the foregoing model.

Darwin's finches on Española

I visited Española in July 1972 and again in January and February 1973. On both occasions I netted, weighed, measured and banded birds in the vicinity of Punta Suárez and searched that area for active nests (unpublished). Some birds were breeding during my first visit, but most breed during the rainy season (January–March)[32]. During my second visit, I collected crop samples from nestlings of both species. On Española, individual *Geospiza fuliginosa* weigh 14–16 g; whereas individual *G. conirostris* weighed from 25 to 39 g. Although these birds differed in body size, their breeding biologies were remarkably similar: both commonly incubated clutches of three eggs (range 1–5) and required about 30 d to complete a clutch and fledge their young; and both showed similar preferences for nest sites (my unpublished results) and the diets of their nestlings were also similar.

The diets of nestling *G. conirostris* contained 17% seeds with the remainder being composed of a variety of invertebrates. *G. fuliginosa* nestlings received a diet composed of 35% seeds. Both species fed the same kinds of seeds to their nestlings (Table 1). Of the over 5,000 seeds recovered from nestling *G. conirostris*, 95% by count and 88% by volume were of one species of grass, *Panicum hirticule*. This species also accounted for more than 86% of the >8,000 seeds recovered from *G. fuliginosa* young, and made up more than 80% of the volume of seeds in those samples. Thus, for both finches, the same grass species accounted for 80–90% of seed material that they fed to their nestlings.

Lepidoptera larvae of various sizes made up over 96% of the

If clutch size is similar for both species, then the larger species will require more food to support its young, since they are larger; and if the breeding season is terminated by a decline in available food, the breeding season of the larger species will end before that of the smaller. A major effect of size will thus be to shorten the breeding season for the larger species (Fig. 2a).

The differences in length of the breeding season for species of different sizes will alter the observed patterns of reproduction at the population level (Fig. 2b and c). Presume first that a few individuals are breeding throughout the year. When a breeding season starts, the establishment of new nests and production of new clutches will shift the relative proportions of nests with eggs away from 50% (Fig. 2c). The deviation of the population towards a surplus of nests with eggs will be countered as soon as the first eggs hatch. The same pattern will occur for the larger species, except that it will occur later. There is thus a period when the smaller species is in the incubation phase (more than half the nests have eggs) of a nesting cycle and the larger species is not (A in Fig. 2c). Since the larger species nested later, its population will seem to remain in incubation phase longer than the smaller species (B in Fig. 2c). As the end of the breeding season approaches, the larger species ceases establishing nests,

Fig. 3 Frequency distribution for the sizes of lepidoptera larvae recovered from nestling *Geospiza fuliginosa* and *G. conirostris* in February 1973. Total no. of larvae: *G. fuliginosa*, 838; *G. coniros-tris*, 172.
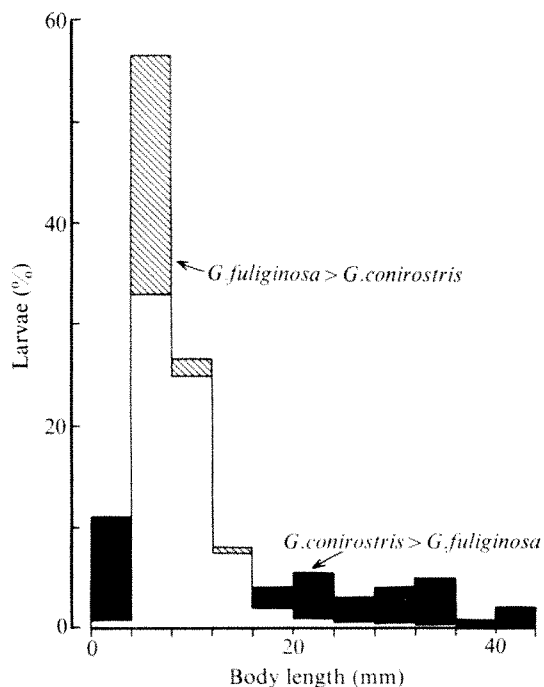
Table 1 The relative numbers of different kinds of seeds recovered from nestling *G. conirostris* and *G. fuliginosa*

Source	Seed volume (mm³)	Percentage occurrence *G. conirostris*	*G. fuliginosa*
Seeds smaller than			
P. hirticule	0.50 ± 0.029	0.6	
	0.13 ± 0.002		1.6
Panicum hirticule	0.69	93.6	86.5
Seeds larger than			
P. hirticule	1.10 ± 0.012	5.6	
	1.06 ± 0.005		11.7
Total number of seeds recovered		5,170	8,262

Table 2 Temporal changes in the diets of nestling *Geospiza conirostris* and *G. fuliginosa*

| Date | Sample size | | Nestling weights (g) | | Animal matter in diet (vol %) | | Samples (%) | | |
| | | | | | | | with large larvae (>20 mm) | with small larvae (20 mm) | |
	con.	ful.	con.	ful.	con.	ful.	con. only	con.	ful.
Feb. 15, 16	5	35	21.2 (1.53)	12.9† (0.26)	95	73	80	20	82
17, 18	26	19	22.2* (0.93)	12.5 (0.35)	91	75	50	42	85
19, 20	18	15	22.1 (0.41)	13.3 (0.41)	72	62	56	61	81
21, 22	7	21	22.1 (0.71)	13.4 (0.28)	76	61	14	43	95
23, 24	10	20	24.3 (0.70)	12.8 (0.33)	85	58	30	80	85
25, 26	8	13	24.4 (0.60)	12.6‡ (0.92)	71	63	25	68	92
27, 28	—	8	—	13.7 (0.62)	—	62	—	—	88
Mar. 1, 2	5	9	18.8 (2.17)	12.3 (0.94)	46	60	0	80	100
Rank correlations with date			−0.22	−0.12	−0.86§	−0.60	−0.86§	0.90¶	0.73§

* $n = 25$; † $n = 14$; ‡ $n = 7$; §$P<0.05$; ¶$P<0.01$

volume of invertebrates recovered from nestling *G. conirostris*, and 76% in the diets of *G. fuliginosa* young (Fig. 3). Small larvae were present in the diets of both species, but large larvae were more common in the diets of *G. conirostris* ($t = 4.04$, $P<0.001$). Over 30% of the larvae recovered from *G. conirostris* nestlings were 20 mm long or longer, whereas larvae of this size accounted for less than 3% of the larvae in the diet of *G. fuliginosa* nestlings. In total, larvae of lepidoptera and the seeds of *P. hirticule* alone accounted for 96% and 80% (by volume) of the food items that *G. conirostris* and *G. fuliginosa* fed to their young.

Nestling diets changed during the course of my study (Table 2). Small larvae (<20 mm in length) always composed 80% or more of the larvae recovered from nestling *G. fuliginosa* during the period February 15 to March 2; whereas, large larvae (>20 mm) became an increasingly rare dietary item for nestling *G. conirostris* and small larvae became more common. This change influenced the overall dietary balance for nestling *G. conirostris*. In mid-February, diets of nestling *G. conirostris* contained 94% animal matter (by volume) whereas by March 2,

animal matter accounted for less than 50% of their diet. Thus, the diets of nestling *G. conirostris* showed a progressive deterioration late in the study, a change that was independent of nestling weight, and not observed in the diets of nestling *G. fuliginosa* (Table 2). These findings are not inconsistent with the assumptions and expectations of the model. The species are similar in a variety of ways, and changes in the nestling diets occur first in the larger species. The temporal patterns of reproduction for both species are also instructive.

It is likely that some breeding occurs throughout the year for both of these species (D. Werner and P. Kramer, personal communication). When I visited Española in January and began to follow the course of activities at various nests, neither species had as yet entered the incubation phase of the breeding cycle (Fig. 4). The *G. fuliginosa* population entered the incubation phase in interval 4 and that population was dominated by nests with eggs until interval 10. The incubation phase began later for *G. conirostris* (interval 8) and a significant majority of nests had eggs until interval 12. Thus, the incubation phase began earlier

Fig. 4 Temporal changes in the percentage of nests with eggs for *Geospiza fuliginosa* (★) and *G. conirostris* (●). An arrow indicates a significant departure from a 1:1 ratio. A double-headed arrow indicates that there are differences between the two species as well. The data are summarised at 2-d intervals beginning on January 22 and ending on March 2, 1973.
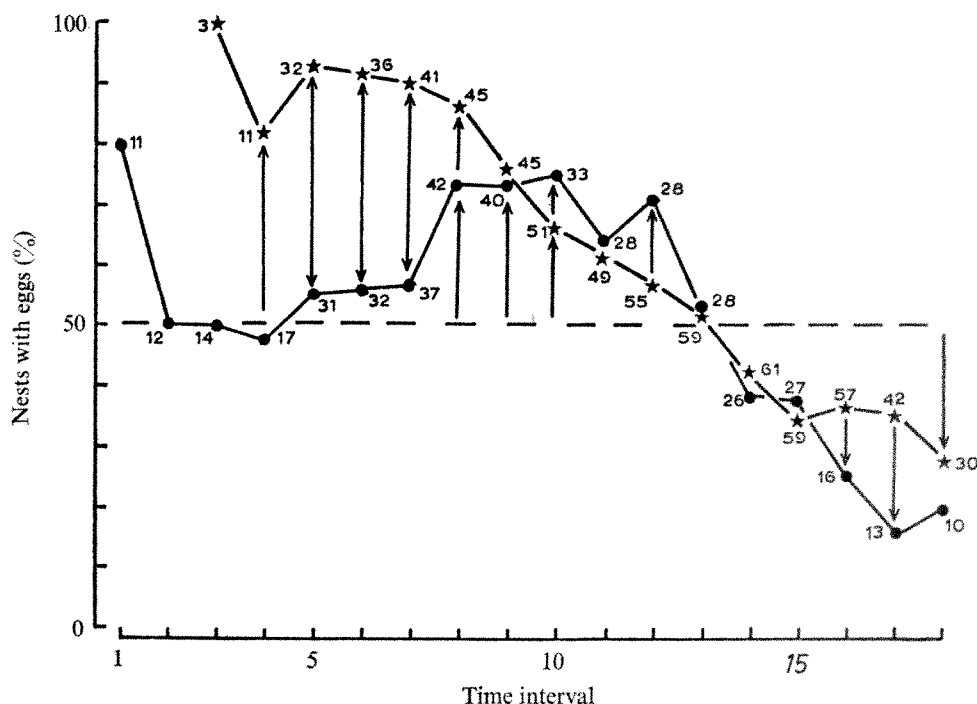
and terminated earlier for *G. fuliginosa*, the smaller species, than it did for *G. conirostris* (compare with Fig. 2c). Exactly the reverse pattern occurred for the onset of the nestling phase. The nestling phase for *G. conirostris* began in interval 16, whereas for *G. fuliginosa* it was delayed until interval 18 (again compare with Fig. 2c). Finally, as expected from the model (Fig. 2b), *G. conirostris* had completed a maximum number of nests (interval 8) before *G. fuliginosa* (interval 14). In short, the observed nesting pattern in both species is identical to that expected from the model and suggests that the larger species, *G. conirostris*, does, indeed have a shorter breeding season than its smaller counterpart.

Seasonal sexual dimorphism in *Geospiza conirostris*

The analysis initially considered comparisons between species because size-related differences in their biology are more easily detected, but the model applies with equal force to individuals of the same species but different size.

During a major breeding season, all females are likely to breed, and so it would be difficult to determine if smaller females of the same species were breeding earlier, later, or were more likely to renest. If a short breeding season occurs, however, then females of small size may breed when the large ones do not.

The Galápagos are affected by the *el niño* phenomenon[33]. This event, which is tied to movements of ocean currents, has the effect of producing wet season (December–March) rains during the dry season (June–August) and thereby creating conditions similar to those in the breeding season. My first visit to Española occurred during an *el niño* year, therefore a comparison of the size of the birds breeding in July with those breeding in January provides a further test of the model.

In July 1972, I netted 225 *G. conirostris*; of which 17 were females with brood patches, and 14 were males in all-black plumage. During the same visit, I located five active *G. conirostris* nests and 15 active *G. fuliginosa* nests. During my winter visit, I netted 231 *G. conirostris*; of which 75 were females with brood patches, and 43 were mature males. In addition, I located 99 *G. conirostris* nests, and the male at 80 of these successfully attracted a female. The occurrence of few active nests in July suggests that *G. conirostris* had recently ceased breeding, and the paucity of adults in reproductive condition indicates that only a few individuals attempted to breed in July.

During the January–March breeding season there was no difference in size between males and females (Table 3). Although the mean size of females was less than that of males, the difference is insignificant. In contrast, however, the females breeding in July were significantly smaller than males both in body weight ($F_{1,30} = 12.5$, $P < 0.001$) and in culmen (beak ridge) length ($F_{1,30} = 5.15$, $P < 0.05$). Further, the females in breeding condition in July were significantly smaller than females that bred in January, both in body weight ($F_{1,91} = 6.78$, $P < 0.01$) and in bill size ($F_{1,91} = 6.60$, $P < 0.05$). Again the observed pattern was as expected from the model.

The extent to which smaller females will have additional opportunities to breed will depend on the frequency with which differential cues occur. The *el niño* vagaries of the Galápagos climate ensure that these events occur with some regularity. A consequence of this phenomenon is to depress the mean size of

females in the population, thus producing a sexually dimorphic population with regard to size.

The seasonal pattern of sexual dimorphism documented for *G. conirostris* and its suggested causes are not limited to large species, rather they should affect all species of finches in the archipelago. This seems to be the case. Lack[17,18] documented, but did not explain, the existence of sexual dimorphism in size in all except one of the Geospizinae—the Cocos Island finch (*Pinaroloxias inornata*), a species restricted to a humid tropical island just off the west coast of Costa Rica.

Size dimorphism in birds

Sexual dimorphism in size may result in differential resource partitioning[1,2,34]. Such partitioning occurs in Darwin's finches[35], but if it were a pervasive selective pressure for sexual dimorphism, it should manifest itself in assortative matings during the breeding season. That is, the male in each pair should consistently be larger, or heavier, than his mate. I have measurements on 12 pairs that bred in January. The male had a larger bill in five instances and was heavier in seven instances. Thus, although the sample is small, it gives no indication that assortative mating was occurring.

Sexual dimorphism has also been attributed to selection for increased size in males, which may be advantageous in male–male competition for breeding sites. In contrast, the finch data indicate that decreased female size resulted in increased fecundity in a fluctuating environment. In particular, small size in females increased their sensitivity to environmental changes that indicate favourable conditions for breeding. Smaller females are thus likely to breed sooner and more often than their larger counterparts.

My inferences concerning sexual dimorphism and its effect on the timing of reproduction in Darwin's finches are based on the consequences of certain bio-energetic relationships and depend on when the resources used for egg production are accumulated. If they are accumulated immediately before reproduction, then relatively small females will be at an advantage because they can accumulate the necessary resources more quickly. For small passerine birds, this is likely to be the case, and in house sparrows (*Passer domesticus*) in particular, Johnston and Sealander[34] have documented a trend of increasing sexual dimorphism with latitude, a finding which is consistent with the model since the timing of reproduction should be more critical in places with short growing seasons. Conversely, if resources for reproduction are derived from reserves stored at some other time, say, before migration, then larger females will be at an advantage because they will have more reserves left for egg production. It is notable that the degree of 'reversed sexual dimorphism' in the three species of Accipiters studied by Storer[13] increases with the degree of migration. Furthermore, certain species of shore birds, in which females are larger than males, arrive on their sub-arctic breeding grounds with substantial fat reserves[37], and the largest females breed first[6]. The timing of reproduction seems to be critical for these species; the arctic growing season is short and early breeding is associated with increased reproductive success[38]. In addition, Brodie[39] has detailed an identical explanation for reversed sexual dimorphism in fin whales (*Balaenoptera physalus*), in which the females store fat while feeding on krill in the Antarctic. They subsequently migrate northwards to their breeding grounds, which are characterised by warmer temperatures and reduced food productivity, where females give birth to and nurse their calves. The energy used during gestation and lactation seems to be derived from fat reserves stored in the Antarctic, and the larger size of females endows them with increased storage capacity to satisfy the demands of reproduction.

Larger male size is characteristic of many species of birds In certain North American species, males do not migrate as far South as females and overwinter in harsher climates, where increased size may reduce overwinter mortality. That migratory pattern seems to ensure the early arrival of males on the breeding ground[40]. In polygynous song birds, males generally arrive well

Table 3 Seasonal sexual dimorphism in *Geospiza conirostris*

Character	Date	Males n	Males Mean	Females n	Females Mean
Body weight (g)	July 1972	14	32.9 (0.66)	17	29.6 (0.62)
	January 1973	43	32.3 (0 37)	75	31.7 (0.34)
Culmen length (mm)	July 1972	14	14.0 (0.21)	17	13.3 (0.24)
	January 1973	43	14.1 (0.12)	75	13.9 (0.10)

Standard errors are given in parentheses.

in advance of the females and establish territories[41-43]. As a consequence, males are subjected to the rigours of territorial establishment at a time when the weather is likely to be inclement. Agonistic encounters are likely to reduce the time available for foraging and thus larger size may provide males with the physiological reserves necessary to hold a territory. This certainly seems to be the case in polygynous mammals. Male northern fur seals (*Callorhinus ursinus*) are 7–10 times the size of females[44], and establish territories before the arrival of the females. Territorial bulls fast for the duration of breeding season (2–4 months)[45]; Bartholomew[46] interprets the extreme sexual dimorphism in this species to be a result of the advantages of early territorial establishment and prolonged territory maintenance, advantages that could be realised by the increased storage capacities associated with increased size.

In summary, simple bioenergetic considerations provide a general description of geographic patterns of size variation, size-related changes in the timing of reproduction and seasonal patterns of sexual dimorphism in Darwin's finches. Expansion of these analyses provides a consistent explanation for patterns of normal and reversed sexual dimorphism in birds and mammals. The analysis accounts for the general condition of larger male size in temperate species and identifies the conditions that would lead to increased female size. In all certainty, these initial tendencies will be amplified by other selective pressures[2].

1 Sealander, R. K., *Condor*, 68, 113–151 (1966).
2 Sealander, R. K., in *Sexual Selection and the Descent of Man* (edit. by Campbell, B.), ch. 8 (Aldine, Chicago, 1972).
3 Pulliam, H. R., *Science*, 189, 474–476 (1975).
4 Denham, W. W., *Am. Anthropologist*, 73, 77–95 (1971).
5 Ingolfsson, A., *Auk*, 86, 732–737 (1969).
6 Jehl, J., *Evolution*, 24, 311–319 (1972).
7 Williamson, P., *Ecol. Monogr.*, 41, 129–152 (1971).
8 Ligon, J. D., *Auk*, 85, 203–215 (1968).
9 Jackson, J. A., *Ecology*, 51, 318–323 (1970).
10 Kilham, L., *Auk*, 87, 544–556 (1970).
11 Wiley, R. H., *Q. Rev. Biol.*, 49, 201–227 (1974).
12 Cade, T. J., *Univ. Calif. Publ. Zool.*, 63, 151–290 (1960).
13 Storer, R. W., *Auk*, 83, 423–436 (1966).
14 Reynolds, R. T., *Condor*, 74, 191–197 (1972).
15 Earhart, C. M., and Johnson, N. K., *Condor*, 72, 251–264 (1970).
16 Schmidt-Nielsen, K., *Science*, 177, 222–228 (1972).
17 Kendeigh, S. C., *Condor*, 72, 60–65 (1970).
18 Lack, D., *Ecological Adaptations for Breeding in Birds* (Methuen, London, 1968).
19 Calder, W., in *Avian Energetics* (edit. by Paynter, R. A., Jr), ch. 3 (Publ. Nuttall Ornith Club, No. 15, Cambridge, 1974).
20 Schaffer, W. M., and Elson, P. F., *Ecology*, 56, 577–590 (1975).
21 Nevo, E., *Ecology*, 54, 1271–1281 (1973).
22 Brown, J. H., and Lee, A. K., *Evolution*, 23, 329–338 (1969).
23 Grubb, T., *Condor*, 77, 175–182 (1975).
24 Bumpus, H. C., *Biol. Bull. Mar. Biol. Lab. Woods Hole*, 1–15.
25 Rising, J. D., in *Productivity, Population Dynamics and Systematics of Granivorous Birds* (edit. by Kendeigh, S. C., and Pinowski, J.) (Warszawa, 1972).
26 Bowman, R. I., *Univ. Calif. Publ. Zool.*, 58, 1–302 (1961).
27 Lack, D., *Darwin's Finches* (Cambridge University Press, London, 1947).
28 Lack, D., *Calif. Acad. Adi., Occas. Papers 21* (1945).
29 Baker, H., *Ecology*, 53, 997–1010 (1972).
30 Schoener, T. W., and Janze, D., *Am. Nat.*, 102, 207–224 (1968).
31 Pianka, E., and Parker, W. S., *Am. Nat.*, 109, 453–464 (1975).
32 Lack, D., *Ibis*, 92, 268–278 (1950).
33 Wyrtki, K., Stroup, E., Patzer, W., Williams, R., and Quinn, W., *Science*, 191, 343–346 (19).
34 Yom-Tov, Y., and Ollason, J. G., *Oikos*, 27, 81–85 (1976).
35 Grant, P. R., Grant, B. R., Smith, J. N. M., Abbott, I. J., and Abbott, L. K., *Proc. natn. Acad. Sci. U.S.A.*, 73, 257–261 (1976).
36 Johnston, R. F., and Sealander, R. K., *Am. Nat.*, 107, 373–390 (1973).
37 Yarburough, C. G., *Auk*, 87, 100–110 (1970).
38 Soikkeli, M., *Ann. Zool. Fennici*, 4, 158–198 (1967).
39 Brodie, P. F., *Ecology*, 56, 152–161 (1975).
40 Ketterson, E. D., thesis, Indiana Univ. (1974).
41 Johm, C. H., *Ecology*, 54, 356–365 (1973).
42 Willson, M. F., *Ecol. Monogr.*, 36, 51–77 (1966).
43 Sealander, R. K., and Giller, D. R., *Condor*, 63, 29–86 (1961).
44 Bartholomew, G. A., *Trans. 18th N. A. Wildlf. Conf.*, 481–502 (1953).
45 Peterson, R. S., in *The Behavior and Physiology of Pinnipeds* (edit. by Rice, C. E.), (Appleton–Century–Crofts, New York, 1968).
46 Bartholomew, G. A., *Evolution*, 24, 546–559 (1970).

Chemical characterisation of the Thy-1 glycoproteins from the membranes of rat thymocytes and brain

A. Neil Barclay, Michelle Letarte-Muirhead* & Alan F. Williams

MRC Immunochemistry Unit, Department of Biochemistry, Oxford OX1 3QU, UK

Ralph A. Faulkes

National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

The Thy-1 antigens from both thymocytes and brain of rats are major membrane glycoproteins of about 25,000 molecular weight of which 30% is carbohydrate. The brain and thymus glycoproteins contain very similar amounts of each amino acid, but have strikingly different carbohydrate compositions. The antigenic determinants are likely to be in the protein part of the molecule.

MANY specific functions of cells are likely to be mediated by tissue specific molecules at the cell surface reacting with other cells or soluble factors. One method of identifying these is to study tissue-specific cell surface antigens, which are commonly referred to as differentiation antigens[1,2]. The theta (θ) antigen, now called Thy-1 (ref. 3) was one of the first cell-surface differentiation antigens to be identified on lymphocytes, and was found in mice in two allelic forms which were serologically identified as the Thy-1.1 (θ-AKR) and Thy-1.2 (θ-C₃H) alloantigenic determinants[4]. The Thy-1 locus is on chromosome 9 in the mouse[5].

The Thy-1 antigen was found in large amounts on mouse thymocytes and brain[4,6] and in smaller amounts on mature T lymphocytes[4,7,8]. The number of antigenic sites per thymocyte was found to be >400,000 for Thy-1.2 (ref. 8) and about 600,000 for Thy-1.1 (ref. 9) by measuring binding of the alloantibodies. The antigenic activity of brain is approximately equal to thymocytes on a packed tissue basis[4]. The Thy-1 antigen has also been reported on mouse epidermal cells and fibroblasts[10,11], but other tissues have very little if any antigen[4]. In the brain the Thy-1 antigen is present in small amounts at birth and increases to maximum specific activity after about 25 d (ref. 6).

The Thy-1.1 antigenic determinant can be recognised on rat thymocytes and brain with mouse alloantiserum, but no rat strain with the Thy-1.2 determinant has been reported[12]. The amount of antigen on rat and mouse thymocytes and brain in similar as is the developmental pattern in the brain but, surprisingly, most peripheral rat T cells lack the antigen, while many rat bone marrow cells express it at the cell surface[2,9,13]. Other tissues of the rat including erythrocytes, heart, kidney, lung and liver have virtually no detectable Thy-1 antigen[2]. Two other antigenic determinants

can be found on the rat Thy-1 molecule and these are recognised by xenogeneic (rabbit) antiserum and are referred to as the rat specific, and the rat–mouse cross reacting, xenoantigenic determinants[2,14]. Similar antigenic determinants have been described in the mouse[15].

There has been considerable confusion as to the chemical nature of the Thy-1 molecule with suggestions that it is a glycolipid[16,17], a protein of various sizes[18,19] or a glycoprotein[20,21] in mouse, and a protein or glycoprotein in rat[22]. Glycoproteins have been purified from rat thymocytes[23] and brain[24] which clearly carry all the Thy-1 antigenic determinants at high specific activities, and the conflicts between these results and the work of others is discussed below. The purifications involved the solubilisation of antigen in deoxycholate from crude membrane preparations, followed by gel filtration and affinity chromatography with columns of lentil lectin coupled to Sepharose-4B. The affinity column bound most of the brain Thy-1, but about 50% of thymocyte Thy-1 did not bind. This fraction was purified by an antibody column against brain Thy-1. Thus three Thy-1 preparations were obtained: brain Thy-1, thymocyte Thy-1 which binds to lentil lectin (Thy-1L+) and thymocyte Thy-1 which does not bind (Thy-1L−). The three preparations could not be distinguished by their antigenic determinants and both thymocyte and brain Thy-1 were of similar immunogenicity in rabbits. In contrast, slight differences in mobility of the glycoproteins were noted after electrophoresis on polyacrylamide gels in sodium dodecyl sulphate. On 12.5% polyacrylamide gels, brain Thy-1 had an apparent molecular weight of 24,000, while for thymocyte Thy-1L+ and Thy-1L− the values were 25,000 and 27,000, respectively[23,24].

We describe here the results of amino acid and carbohydrate analysis of the Thy-1 glycoproteins and also report experiments to determine the nature of the antigenic determinants.

Preparation of Thy-1 samples for chemical analysis

The three forms of Thy-1 were purified from rat thymocytes and brain as before[23,24] (except that twice recrystallised sodium deoxycholate was used) and the minor impurities in thymocyte Thy-1L− were removed by gel filtration on a column of Sephadex G-200 in 0.5% sodium deoxycholate. The purified fractions contained a large amount of deoxycholate after concentration by ultrafiltration, and this was removed by precipitation of the glycoprotein with ethanol in which deoxycholate is soluble (M. J. Crumpton, personal communication). Ethanol was added to a final concentration of 75%, together with a few drops of saturated sodium acetate in ethanol (to improve flocculation of the precipitate) and the mixture was left for 48 h at −20 °C. The precipitate was recovered and washed in absolute ethanol by centrifugation at 1,800 g for 30 min at −10 °C. The precipitates were readily soluble in water, and a small amount of turbidity was removed by centrifugation at 100,000g for 30 min. The resulting pellets contained less than 5% of the Thy-1 by antigenic activity or amino acid and carbohydrate analysis, and probably consisted of trace impurities from the deoxycholate.

This precipitation gave yields of 70% for both thymocyte forms and 40% for Thy-1 from brain. The procedure had no effect on the antigenicity of the glycoproteins whose specific activities for Thy-1.1 and Thy-1 xenoantigenic determinants were similar to those of the pure preparations before precipitation[23,24]. The electrophoretic behaviour of the glycoproteins on polyacrylamide gels in sodium dodecyl sulphate (SDS) was also unaffected by the precipitation, and, as Fig. 1 shows, none of the three preparations showed detectable contaminant bands when stained for protein or carbohydrate.

Preliminary studies on the size of Thy-1 in the absence
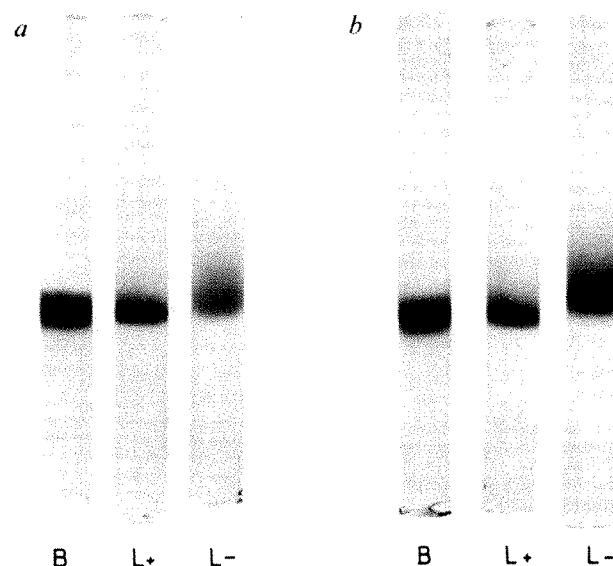
Fig. 1 Purity of Thy-1 fractions analysed by polyacrylamide gel electrophoresis in sodium dodecyl sulphate. B, L+ and L− denote Thy-1 purified from brain, and Thy-1L+ and Thy-1L− from thymocytes respectively, staining for protein with Coomassie blue (*a*) or for carbohydrate with periodic acid–Schiff (*b*) stain[23]. The Thy-1 glycoproteins which had been precipitated with ethanol and solubilised in water, were electrophoresed on 7.5% polyacrylamide gels in SDS[23]. 14 µg and 22 µg of the glycoproteins were loaded for gels stained for protein and carbohydrate, respectively.

of deoxycholate have been carried out by sedimentation equilibrium methods in the ultracentrifuge, and most Thy-1 was in an aggregate of 250,000–350,000 weight-average molecular weight (unpublished results of P. Kuchel). After addition of deoxycholate, Thy-1 was found in an antigen–detergent complex of molecular weight 29,000, in agreement with previous studies[22,24]. Aggregation in the absence of detergent would be expected for a membrane molecule which was normally associated with lipids.

Table 1 Amino acid composition of Thy-1 glycoproteins

| | No. of amino acids per 100 residues | | |
| | Brain | Thymocyte | |
	Thy-1	Thy-1L+	Thy-1L−
Asx	12.7	12.6	13.0
Glx	9.1	9.1	9.1
His	4.1	3.8	4.0
Lys	6.9	7.2	7.0
Arg	7.5	7.5	7.2
Thr	7.6	8.6	8.3
Ser	7.4	7.1	7.9
Pro	3.6	3.8	3.0
Ala	3.3	2.9	3.2
Cys	3.1	3.2	3.2
Gly	6.0	4.9	5.0
Tyr	2.0	2.0	2.0
Val	7.3	7.4	7.4
Ile	3.9	4.1	4.0
Leu	10.4	11.2	11.1
Phe	4.0	3.7	3.8
Met	1.1	0.9	0.8

Amino acid analyses are expressed as the mean number of each amino acid residue per 100 residues and are the result of at least four analyses in each case. For brain Thy-1 and thymocyte Thy-1L+, two completely independent preparations of each were analysed, while for thymocyte Thy-1L− one preparation was used. The amino acids are arranged into hydrophilic, intermediate and hydrophobic groups[25] reading down the table. Samples for amino acid analyses were hydrolysed with constant boiling HCl at 110 °C for 20 h and analysed on a Locarte or modified Beckman analyser[26]. A correction for the destruction of threonine and serine was calculated from 24-and 72-h hydrolyses of Thy-1 from brain and applied to all analyses. Cysteine was estimated as cysteic acid after performic oxidation[27]. The s.e.m. was < 5% of each mean value except for proline values (< 12%).

Chemical analysis of the Thy-1 glycoproteins

The preparations of Thy-1 shown in Fig. 1 were analysed for their amino acid (Table 1) and carbohydrate compositions (Table 2) by standard procedures. The amino acid analyses were very similar for the three forms of Thy-1 and the only differences which were significant after statistical analysis (not shown) were those between the leucine and glycine content of brain Thy-1 in comparison with thymocyte Thy-1. The compositions were so similar, however, that any differences which may exist in the protein part of the molecules will only be clearly revealed by amino acid sequence analysis. If the amino acids are grouped on the basis of their hydrophobicity (Table 1), it can be seen that Thy-1 is much less hydrophobic than some membrane proteins[24] and indeed is notable for its large proportion of hydrophilic residues. Hydrophobicity in amino acid compositions is, however, not an invariable feature of membrane proteins and the major human red cell glycoprotein which is superficially similar to Thy-1 is also quite hydrophilic in its amino acid composition[27]. In detail the amino acid compositions of the red cell glycoprotein and Thy-1 are very dissimilar.

If the molecular weight of 25,000 as determined by electrophoresis on polyacrylamide gels in SDS is taken as correct, then the size of the polypeptide portion would be about 17,000 and the values in Table 1 should be multiplied by 1.5 to give the number of residues per molecule.

In contrast to the amino acid analyses, marked differences in carbohydrate composition were found particularly between brain and thymocyte Thy-1 with smaller differences between thymocyte Thy-1L+ and Thy-1L−. These results are shown in Table 2 and are expressed as residues per 100 residues of amino acids. If the figures are multiplied by 1.5, an estimate of residues per molecule can be obtained. Galactosamine was found only in brain samples and no trace was detected in thymocyte Thy-1. In contrast, sialic acid was present in very small amounts in brain Thy-1, but in much larger amounts in thymocyte Thy-1. The amounts of fucose, glucose and galactose also differed by a factor of 2 or more between brain and thymocyte Thy-1, while smaller differences on a percentage basis were found between mannose and glucosamine contents. Given that differences in composition are likely to reflect much larger differences in structure, the carbohydrate chains of brain and thymocyte Thy-1 may be completely unrelated.

Between thymocyte Thy-1L+ and Thy-1L− the compositions were much more similar, but differences in the amount of galactose, mannose, glucosamine and sialic acid may be significant. These differences in composition do not provide a simple answer as to why Thy-1L− fails to be bound by the lentil lectin, except to reinforce the view that it is likely to be due to differences in the carbohydrate rather than the polypeptide structure. The differences between Thy-1L+ and Thy-1L− are typical of the micro-heterogeneity of carbohydrate residues observed for many glycoproteins[23]. The large differences between brain Thy-1 and thymocyte Thy-1, however, represent a much more unusual order of heterogeneity of carbohydrates and could reflect a similar polypeptide chain carrying unrelated carbohydrate structures in the two tissues.

In the carbohydrate analysis by gas–liquid chromatography only small amounts of material other than known sugars were found; these included traces of lipid (see legend to Table 2). The procedures used in the purification of Thy-1 involving deoxycholate and ethanol would be expected to remove non-covalently-bound lipids and previously no lipid-bound phosphorus (less than 0.05 mg of phospholipid per mg of Thy-1) had been detected in brain Thy-1 (ref. 24).

From a summation of the amino acid and carbohydrate analyses, the percentage composition by weight of the molecules was calculated (Table 2). The values for carbo-

hydrate content of about 30% are higher than for many membrane glycoproteins.

Chemical basis of Thy-1 antigenic determinants

Rat Thy-1, regardless of its tissue origin, carries at least three sets of antigenic determinants. These include Thy-1.1 which is recognised by mouse alloantiserum[8,13] and two xenoantigenic determinants recognised by rabbit serum[9,14]. The similarity between the amino acid compositions, and the differences in carbohydrate compositions of Thy-1 from brain and thymocytes suggest that these antigenic determinants are defined by the protein rather than the carbohydrate structure. To investigate this point further the susceptibility of the antigenicity to heating and proteolysis was studied. Both of these procedures should destroy protein determinants which are usually dependent on conformation[34] but leave carbohydrate antigenic determinants intact[35].

Samples of pure Thy-1 were heated for 10 min at various temperatures and assayed for Thy-1.1 and Thy-1 xenoantigenic determinants (Thy-1 rat specific and rat–mouse cross reacting xenoantigenic determinants were not distinguished). Figure 2 shows a typical experiment in which the activity of these antigenic determinants of brain Thy-1 was destroyed in the temperature range 70–80 °C. Similar results were obtained in experiments with thymocyte Thy-1L+.

The effects of proteolytic digestion were evaluated by measuring the loss of antigenic activity with the inhibition assays, and destruction of the polypeptide chain by SDS–polyacrylamide gel electrophoresis. Thy-1 was quite resistant to all the proteolytic enzymes tested, but after incubation

Temperature (°C) of preincubation
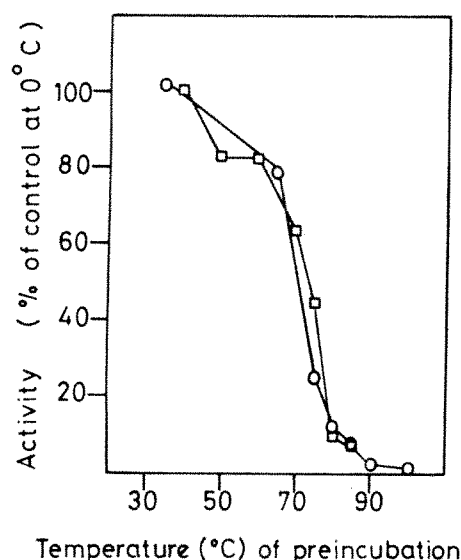
Fig. 2 Effect of heat on the Thy-1 antigenic determinants. The figure shows the results of two experiments on heating of brain Thy-1, assaying for Thy-1.1 (○) and Thy-1 xenoantigenic activities (□). Purified Thy-1 (approximately 10 µg ml⁻¹) was heated for 10 min at the temperature shown in 0.1 % sodium deoxycholate, 0.25 % bovine serum albumin, 0.02 % NaN₃ and 0.01 M Tris-HCl, pH 8.0. The samples were cooled to 0 °C and assayed for Thy-1.1 and Thy-1 xenoantigenic activities[23]. These activities are expressed as the percentage activity compared with a similar sample kept at 0 °C throughout (100 %).

with Pronase for 24 h at 37 °C, 80–95% of the antigenic activity of all Thy-1 determinants was destroyed (**Table 3**). Also the polypeptide chain was almost completely degraded as shown for brain Thy-1 and thymocyte Thy-1L+ in Fig. 3a and b, respectively. Pronase digestion for 3 h caused only partial loss of antigenic activity and partial removal of the polypeptide chain. When brain or thymocyte Thy-1 was incubated for 24 h with trypsin, chymotrypsin or papain (conditions as for Pronase in Table 3 except for the addition of cysteine HCl with papain), there was only partial loss (30–80%) of antigenic activity and the Thy-1 molecule. On a qualitative basis no discrepancies were observed between the amount of antigenic activity and the amount of Thy-1 band remaining on the polyacrylamide gel.

These results, together with the chemical analysis (Tables 1 and 2), and the experiments on the effect of heating, strongly (although not conclusively) suggest that all the antigenic determinants so far identified on rat Thy-1 are expressed by the protein part of the molecule.

Confusion about the nature of the Thy-1 molecule

There has been considerable disagreement about the chemical nature of the Thy-1 molecule. Most other workers have studied mouse Thy-1 but the basic characteristics of the molecule are likely to be similar in both mouse and rat.

Table 3 Effect of Pronase on Thy-1 antigenic activities

| | Residual antigenic activities (%) | |
	Thy-1.1 antigenic activity	Thy-1 xenoantigenic activity
Thy-1 (brain)	3, 7, 6	5, 6, 4
Thy-1L+ (thymocyte)	16, 10, 11	14, 9, 8

Purified brain Thy-1 or thymocyte Thy-1L+ was incubated at 37 °C for 24 h in the presence or absence of Pronase. The antigenic activities of the digested and control preparations were then assayed and are shown for three independent experiments as the percentage of the digested material compared with the control incubation. In the incubation Thy-1 was at 400 µg ml⁻¹ in 0.25 % Na deoxycholate, 0.02 % NaN₃ and 0.01 M Tris-HCl, pH 8.0, and Pronase was added at a weight ratio of 1 : 12 compared with antigen. The assay of antigenic activity was done at 0 °C in the presence of a 1,500-fold excess of bovine serum albumin to inhibit residual Pronase activity.

A glycoprotein band similar to rat Thy-1 can be identified in mouse thymocyte membrane, and the mouse Thy-1.1 antigenic activity is found in the same place as rat Thy-1 after gel filtration in deoxycholate[23]. Also Trowbridge et al., using hyperimmune serum to rat brain Thy-1, could immunoprecipitate a glycoprotein of molecular weight about 25,000 from mouse thymocytes which had been surface labelled with ¹²⁵I and solubilised in detergent[20]. This molecule was present in tissue culture cell lines which displayed Thy-1 antigen, but absent from mutants which had become negative for the antigen[21]. Using conventional anti-(Thy-1) alloantiserum they could not precipitate the 25,000 molecular weight glycoprotein and this confirmed the observations of Vitetta et al.[16]. The latter authors interpreted this to mean that Thy-1 antigens were inactivated by detergent, but this was clearly shown to be incorrect in studies using inhibition of radioimmunoassays[22]. It is more likely that the failure to precipitate Thy-1 was due to low affinity of conventional alloantiserum, leading to dissociation of antigen–antibody complexes in the washing of the immunoprecipitates. In similar experiments Atwell et al.[18] immunoprecipitated a protein of apparent molecular weight 60,000. This result has not been confirmed by others, and may have been due to the use of nonspecific antiserum.

Sauser et al. have suggested that the Thy-1.2 alloantigen is associated with a protein of molecular weight about 40,000 which they obtained in preparations purified 60-fold for the antigen[19]. Thy-1 glycoprotein in the rat was pure when Thy-1.1 antigenic activity was purified 260-fold[23] and thus it is possible that the protein of Sauser et al.[19] is a major contaminant in a fraction which also contains Thy-1 glycoprotein.

Miller and Esselman have suggested that mouse Thy-1 antigen is a ganglioside[17] and this is completely inconsistent with the antigenic determinants being in the protein part of a glycoprotein. Their inhibitions of Thy-1 cytotoxicity assays with gangliosides, however, do not show the high degree of specificity expected for the Thy-1 alloantigens as defined by Reif and Allen[4]. Thus Miller and Esselman suggest that GM₁ ganglioside carries the Thy-1.2 determinant, yet ganglioside from a Thy-1.1 strain inhibits a

Fig. 3 Effect of Pronase on the Thy-1 glycoproteins. The Thy-1 glycoproteins were digested with Pronase as described in Table 3 and the effects analysed by electrophoresis on 7.5 % polyacrylamide gels in SDS (as in Fig. 1). Equal amounts of protein from each digest and its control (about 5 µg) were electrophoresed and after staining for protein with Coomassie blue the gels were scanned at 550 nm[23]. a, Scans for brain Thy-1 control (—) and brain Thy-1 after Pronase digestion (- - -). b, Thymocyte Thy-1L+ control (—) and Thy-1L+ after Pronase digestion (- - -).
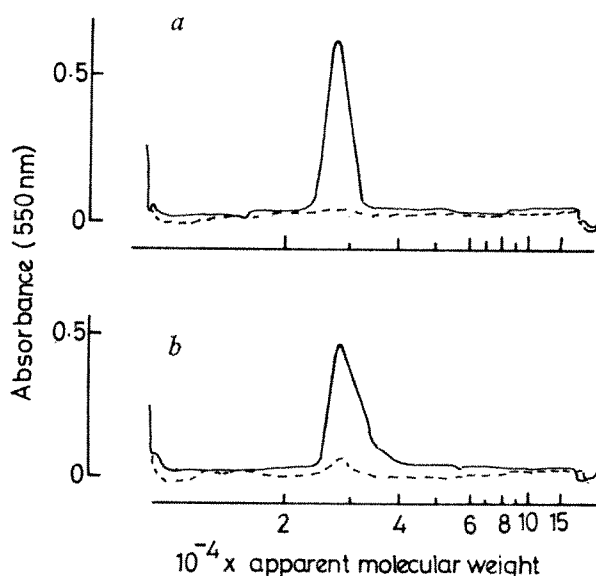
cytotoxicity assay for Thy-1.2, 50% as well as ganglioside from a Thy-1.2 mouse[17]. Also, very large amounts of ganglioside were needed for inhibition of the cytotoxicity assay for Thy-1.2, for example for 50% inhibition approximately 500 ng of GM$_1$ (molecular weight 1,500) was required, compared with 10 ng of Thy-1 glycoprotein (molecular weight 25,000) for 50% inhibition of radioactive binding assays for Thy-1.1[13,14]. Finally, other authors have not been able to find Thy-1 antigenic activity in lipid fractions[24,37].

Trowbridge *et al.* agree that Thy-1 is carried by a glycoprotein[20], but have suggested that the antigenic determinants of mouse Thy-1 are on the carbohydrate and not the protein part of the molecule[21]. This view was based on studies on mutant cell lines, but there is no obvious reason why the data cannot be interpreted equally well with the antigenic determinants carried by the protein part of the molecule.

Possible function for Thy-1 in thymus and brain

In mouse, peripheral T cells have Thy-1 and bone marrow cells do not[7], whereas in the rat the converse is true[11]. Thus it is possible that Thy-1 is not of functional significance in these cell types of either species. In thymus and brain, Thy-1 is a major membrane glycoprotein and these are likely to be the tissues in which the Thy-1 molecules have a biological function.

A large part of the protein of the Thy-1 molecule is likely to be exposed at the cell surface since its amino acid composition is hydrophilic and the antigenic determinants are clearly available on intact cells. By analogy with other membranes, the carbohydrate structures are also likely to be exposed[33]. Although there is no evidence about the function of Thy-1, the molecular properties do provide some limitations. For example, if the polypeptide chain mediates the specific function of Thy-1, the functions are presumably the same in both thymus and brain. Alternatively, the specific function may be mediated by the carbohydrate structures. In this case the polypeptide chain may anchor the molecule in the membrane and provide a backbone for the display of different carbohydrate ligands in thymus and brain. Cell surface carbohydrate has been suggested to be involved in cell–cell interactions by providing recognition sites for glycosyl transferases or lectins on other cells[38–40]. These ideas have gained particular emphasis in studies on slime moulds where strain specific lectins have been identified. These are found at the cell surface and may be involved in the aggregation of single cells in the slime mould's life cycle[40,41]. In thymocytes and brain, Thy-1 would be an obvious possibility as a molecule involved in functions of this type. If this is the case, the differences in carbohydrate composition between brain and thymus Thy-1 may reflect differences in the specificities of recognition in the two tissues.

We thank Mr. T. Gascoyne for amino acid analyses; Ms Sheila Lathwell for amino sugar analyses and Ms Marilyn Simpkins for technical assistance. Drs. M. J. Crumpton and D. Snary gave advice on the precipitation of glycoproteins and analysis of carbohydrate. M. L-M. was a centennial fellow of the MRC of Canada and A.N.B. was supported by a UK MRC research fellowship.

1 Bennett, D., Boyse, E. A., and Old, L. J., in *Cell Interactions*, Third Lepetit Colloquium (edit. by Silvestri, L. G.), 247–263 (North-Holland, Amsterdam, 1972).
2 Morris, R. J., and Williams, A. F., *Eur. J. Immun.*, 5, 274–281 (1975)
3 Snell, G. D., and Cherry, M., in *RNA Viruses and Host Genome in Oncogenesis* (edit. by Emmelot, P., and Bentvelzen, P.), 221–228 (North Holland, Amsterdam, 1972).
4 Reif, A. E., and Allen, J. M. V., *J. exp. Med.*, 120, 413–433 (1964).
5 Blankenhorn, E P., and Douglas, T. C., *J Hered.*, 63, 259–263 (1972).
6 Reif, A. E., and Allen, J M. V., *Nature*, 209, 523 (1966)
7 Raff, M. C', *Transplant. Rev*, 6, 52–80 (1971).
8 Hammerling, U., and Eggers, H. J., *Eur J Biochem.*, 17, 95–99 (1970)
9 Acton, R. T, Morris, R. J., and Williams, A. F., *Eur. J Immun.*, 4, 598–602 (1974)
10 Schend, M., Boyse, E. A., Carswell, E. A., and Old, L. J., *J. exp. Med.*, 135, 938–955 (1972).
11 Stern, P L, *Nature new Biol.*, 246, 76–78 (1973)
12 Douglas, T. C., *J. exp. Med.*, 136, 1054–1062 (1972).
13 Williams, A F., *Eur J. Immun.* 6, 526–528 (1976).
14 Morris, R. J., Letarte-Murhead, M., and Williams, A F., *Eur. J. Immun.*, 5, 282–285 (1975).
15 Thiele, H-G., and Stark, R., *Immunology*, 27, 807–813 (1974).
16 Vitetta, E. S., Boyse, E. A., and Uhr, J W., *Eur J. Immun.*, 3, 446–453 (1973).
17 Miller, H C., and Esselman, W. J, *J. Immun.*, 115, 839–843 (1975)
18 Atwell, J. L., Cone, R E., and Marchalonis, J. J., *Nature new Biol.*, 241, 251–252 (1973).
19 Saunor, D., Anckern, C, and Bron, C, *J Immun.*, 113, 617–624 (1974).
20 Trowbridge, I. S., Weissman, I. L, and Bevan, M. J., *Nature*, 256, 652–654 (1975)
21 Trowbridge, I. S, and Hyman, R., *Cell*, 6, 279–287 (1975).
22 Letarte-Murhead, M., Acton, R T, and Williams, A. F., *Biochem J.*, 143, 51–61 (1974)
23 Letarte-Murhead, M., Barclay, A. N, and Williams, A. F, *Biochem. J.*, 151, 685–697 (1975)
24 Barclay, A. N, Letarte-Murhead, M., and Williams, A F., *Biochem J*, 151, 699–706 (1975).
25 Stock, T. L, and Fox, C. F., in *Membrane Molecular Biology* (edit. by Fox, C F, and Keith, A. D.), 27–75 (Sinauer Associates, Stamford, Connecticut, 1972).
26 Reid, K. B. M., *Biochem. J.*, 141, 189–203 (1974).
27 Hirs, C. H. W., in *Methods in Enzymology*, 11 (edit. by Hirs, C. H W), 59–62 (Academic, New York, 1967).
28 Liu, T.-Y, and Chang, Y. H, *J. biol. Chem*, 246, 2842–2848 (1971).
29 Clamp, J. R., *Biochem. Soc. Symp*, 40, 3–16 (1974)
30 Snary, D., Allen, A. K., Faulkes, R A., and Crumpton, M. J, *Biochem. J*, 153, 75–78 (1976).
31 Hess, H. H., and Rolde, E., *J. biol. Chem*, 239, 3215–3220 (1964)
32 Marchesi, V. T, Tillack, T. W., Jackson, R. L., Segrest, J. P., and Scott, R. E., *Proc. natn. Acad. Sci U S A.*, 69, 1445–1449 (1972)
33 Montgomery, R., in *Glycoproteins* (edit. by Gottschalk, A.), 518–528 (Elsevier, Amsterdam, 1972)
34 Crumpton, M J., in *Antibodies to Biologically Active Molecules* (edit. by Cinader, B.), 61–84 (Pergamon, Oxford, 1967).
35 Watkins, W. M, in *Glycoproteins* (edit. by Gottschalk, A.), 830–891 (Elsevier, Amsterdam, 1972).
36 Arndt, R., Stark, R., Klein, P., Muller, A., and Thiele, H.-G, *Eur J. Immun.*, 6, 333–340 (1976).
37 Kucich, U. N., Bennett, J. C, and Johnson, B J., *J Immun*, 115, 626–630 (1975)
38 Hughes, R. C., *Essays in Biochem*, 11, 1–36 (1975).
39 Roth, S., McGuire E. J., and Roseman, S., *J. Cell Biol*, 51, 536–547 (1971).
40 Rosen, S D., Kafka, J. A., Simpson, D L., and Barondes, S. H, *Proc natn Acad Sci A.*, 70, 2554–2557 (1973)
41 Chang, C.-M., Reitherman, R. W., Rosen, S. D., and Barondes, S H., *Expl Cell Res*, 95, 136–142 (1975)

Nucleotide rigidity

Frederick E. Evans & Ramaswamy H. Sarma

Department of Chemistry, State University of New York at Albany, Albany, New York 12222

It is shown that in aqueous solution the backbone conformation of adenosine is as much flexible as that of 3'-AMP, 5'-AMP and 3', 5'-ADP indicating that nucleotides are not any more rigid than nucleosides. The flexible conformation of the monomeric components is conserved in the nucleotidyl units of destacked ApA, ApApA and poly(A), but it is not conserved in base stacked conditions. The findings are extended to guanosine, uridine and cytidine systems. It is projected that in aqueous solution, conformations of the individual nucleotidyl units of yeast tRNA^{Phe} are confined to the classically stable domains in the base stacked region and non-rigid flexible structures populate in the unstacked region comprising D16, D17, G20, U47 and A76.

MANY investigations have been carried out on the conformation of mononucleotides in order to understand better the conformation of biologically functional nucleic acids. It has been proposed that the conformation of mononucleotides is rigid, and because of this their conformation is conserved in polynucleotides[1–3]. The essential elements of the concept of a conforma-
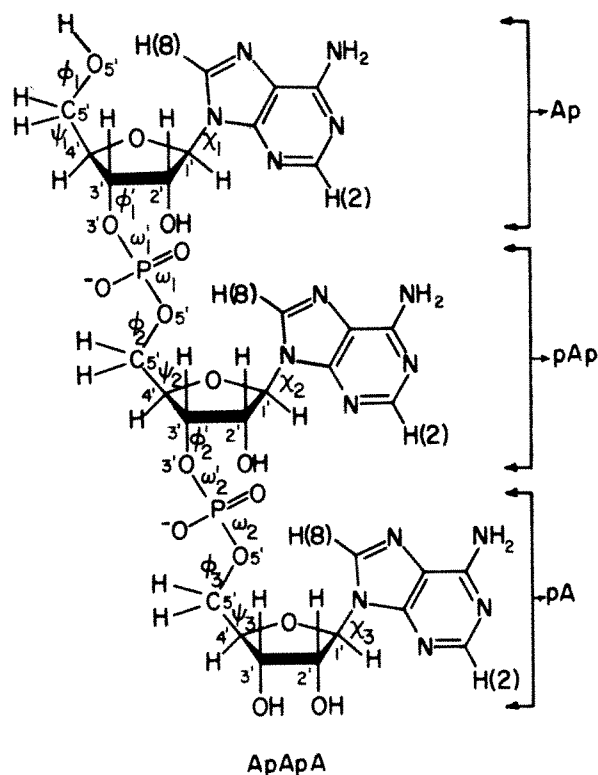
ApApA

Fig. 1 Structure of ApApA, showing the atom numbering scheme and notations for the torsion angles (IUPAC-IUB recommendations).

tionally rigid nucleotide[1] are (1) that the nucleotides are conformationally more rigid than the corresponding nucleosides and that the phosphate group imparts rigidity to the nucleotides (this has been termed the 'phosphate effect'[1,2]) and (2) that polymers achieve their biologically functional structure by rotation around the phosphodiester bonds. Quantum mechanical calculations suggested that the intrinsic conformational rigidity of a nucleotide may not be fundamentally different from that of the corresponding nucleoside[3]. The concept of rigidity is based on the finding that in the solid state nucleosides exhibit a constant conformational pattern—glycosidic torsion = *anti*; sugar pucker = ^3E; $\psi \simeq 60°$ and $\phi \simeq 180°$—which is absent in nucleotides[1,2]. Recent X-ray crystallographic studies have found several nucleotides with conformations different from that of a rigid nucleotide[4-9]. Some differences have been found between solid state and solution state conformations[10-12]. We have undertaken a systematic study of the fine conformational features in aqueous solution of adenosine (A), 3'-AMP

(Ap), 5'-AMP (pA), 3',5'-ADP (pAp), ApA, ApApA (Fig. 1) and poly(A) as well as uridine, 3'-UMP (Up), 5'-UMP (pU), UpU and poly(U) using high frequency nuclear magnetic resonance (NMR) spectroscopy combined with a Karplus-type analysis of the coupling constants. Here we examine the results obtained with respect to the rigid nucleotide concept.

All materials used were commercial preparations. ¹H NMR spectra were recorded at 100, 270 or 300 MHz in the Fourier transform mode using systems described elsewhere[13-15]. All NMR spectra were analysed using a UNIVAC 1108 computer and LACON III. The observed and simulated spectra of ApApA is shown in Fig. 2 and the data are summarised in Table 1.

Conformation in destacked oligomer and polymer

The backbone linkage is comprised of the C4'-C5'-O5'-P-O3'-C3' bond network. The conformation of the C4'-C5' and C5'-O5' bonds shows itself in the coupling constant sums $\Sigma(J_{H4'-H5'} + J_{H4'-H5''})$ and $\Sigma'(J_{H5'-P}-J_{H5''-5})$ respectively[18]. The near identity in the values of Σ(Table 1, Fig. 2) for each of the adenylyl parts in mostly destacked ApApA (72 °C) clearly shows the strong similarity in the time-average conformation about the three C4'-C5' bonds. The Σ data in Table 1 further reveal that this C4'-C5' bond conformation is very similar to that in mostly destacked ApA (72 °C) as well as to that in 3',5'-ADP, 3'-AMP, 5'-AMP and adenosine. Likewise, the Σ' data in Table 1 reveal that the conformational preference about the C5'-O5' bonds in destacked ApApA and ApA is strongly similar to that of the same bonds in 5'-AMP and 3',5'-ADP. The observed values of Σ and Σ' can be translated into an estimate of the percentage population of conformers using empirical equations developed in this laboratory[18]. These equations have been recently modified[19]. Computation of the population distribution of conformers shows that the *gauche-gauche* (ψ, ψ_1, ψ_2, $\psi_3 = 60°$) population is 70–80%, and the combined population of *gauche-trans* (ψ, ψ_1 ψ_2, $\psi_3 = 180°$) and *trans-gauche* (ψ, ψ_1, ψ_2, $\psi_3 = 300°$) conformers is 20–30% irrespective of whether one is dealing with a nucleoside, 3'-nucleotide, 5'-nucleotide, 3'5'-diphosphonucleoside or a destacked dinucleoside monophosphate or a trinucleoside diphosphate. The *gauche'-gauche'* (ϕ, ϕ_1 ϕ_2 $\phi_3 = 180°$) population about C5'-O5' for all these compounds lies between 65–75%. The observed time-average preference for the *gauche-gauche* and *gauche'-gauche'* arrangement is further substantiated by the 1.5–1.8-Hz magnitude of the four-bond $^4J_{H4'-P5'}$ coupling [20].

Information on the conformation about the C3'-O3' part of the backbone can be obtained from the magnitude of the coupling constant $J_{H3'-P3'}$ (refs 15 and 21). The observed

Table 1 Coupling constant data for the various nucleic acid systems of the adenosine family

Temperature (°C) Solvent	Adenosine 22 D₂O	3'-AMP 22 D₂O	5'-AMP 22 D₂O	5'-AMP 86 DMSO+D₂O	3',5'-ADP 22 D₂O	ApA Ap- 12 D₂O	ApA -pA 12 D₂O	ApA Ap- 72 D₂O	ApA -pA 72 D₂O	ApApA* Ap- 18 D₂O	ApApA* pAp 18 D₂O	ApApA* -pA 18 D₂O	ApApA* Ap- 72 D₂O	ApApA* pAp 72 D₂O	ApApA* -pA 72 D₂O	poly(A) 22 D₂O	poly(A)† 86 DMSO+D₂O
H1'-H2'	6.2	6.2	5.7	5.8	5.5	3.6	4.0	4.8	4.6	3.4	2.8	3.5	4.9	5.0	4.8	2	5.9
H2'-H3'	5.2	5.2	5.2	5.2	4.7	5.2	5.3	5.2	5.1				5.2	5.2	5.2		5.2
H3'-H4'	3.5	2.9	3.7	3.5	3.7	5.5	5.8	4.4	4.6				(4.4)‡	4.7	4.9		3.4
H3'-P3'	–	7.9	–	–	8.5	7.6	–	8.8	–				(7.5)‡	7.9	–		8.0
H4'-H5'	2.8	2.4	3.1	3.3	2.9	2.5	2.8	2.6	2.8				2.8	2.7	2.8		3.5
H4'-H5"	3.6	3.3	3.1	3.9	2.9	3.5	3.7	4.0	4.2				3.9	3.8	3.8		3.8
Σ	6.4	5.7	6.2	7.2	5.8	6.0	6.5	6.6	7.0				6.7	6.5	6.6		7.3
H4'-P5'	–	–	1.7	1.5	2.2	–	2.0	–	1.8				–	1.8	1.8		1.7
H5'-P5'	–	–	5.0	5.7	5.3	–	3.0	–	4.9				–	-4.8	5.0		5.7
H5"-P5'	–	–	5.0	5.6	5.3	–	3.3	–	4.9				–	4.6	5.0		5.6
Σ'	–	–	10.0	11.3	10.6	–	6.3	–	9.8					9.4	10.0		11.3
H5'-H5"	–12.6	–11.7	–12.0	–12.0	–11.6	–13.0	–12.0	–12.8	–11.6				–12.8	–11.8	–12.0		–12.0

Concentration is 0.02 M for ApA and ApApA, pD 7.0. Concentration is 0.05 M for the mononucleotides, pD 5.4. The pD is such that the phosphate will be mostly monoanionic like the oligomers and polymers. Data for the monomers at 22 °C can be compared with data for oligomers and polymers at higher temperatures because at the concentration used, the monomer conformations are not significantly perturbed in the temperature range. See, for example, data for 5'-AMP at 22 °C in D₂O and at 86 °C in DMSO-D₂O. Error in coupling constants ±0.1 Hz unless otherwise stated.
* The assignments for the three 1' resonances proposed by Kondo *et al.*[16] has been verified by the incremental methods of Tso' *et al.*[17] by comparing ApA, ApApA and ApApApA.
† Poly(A) dissolved in 50% d₆-DMSO and 50% D₂O at 86 °C, 0.05 M in mononucleotide. For error in the analyses see legend of Fig. 4.
‡ These values are approximate because of the presence of extensive second order perturbations at the frequency used of 270 MHz.
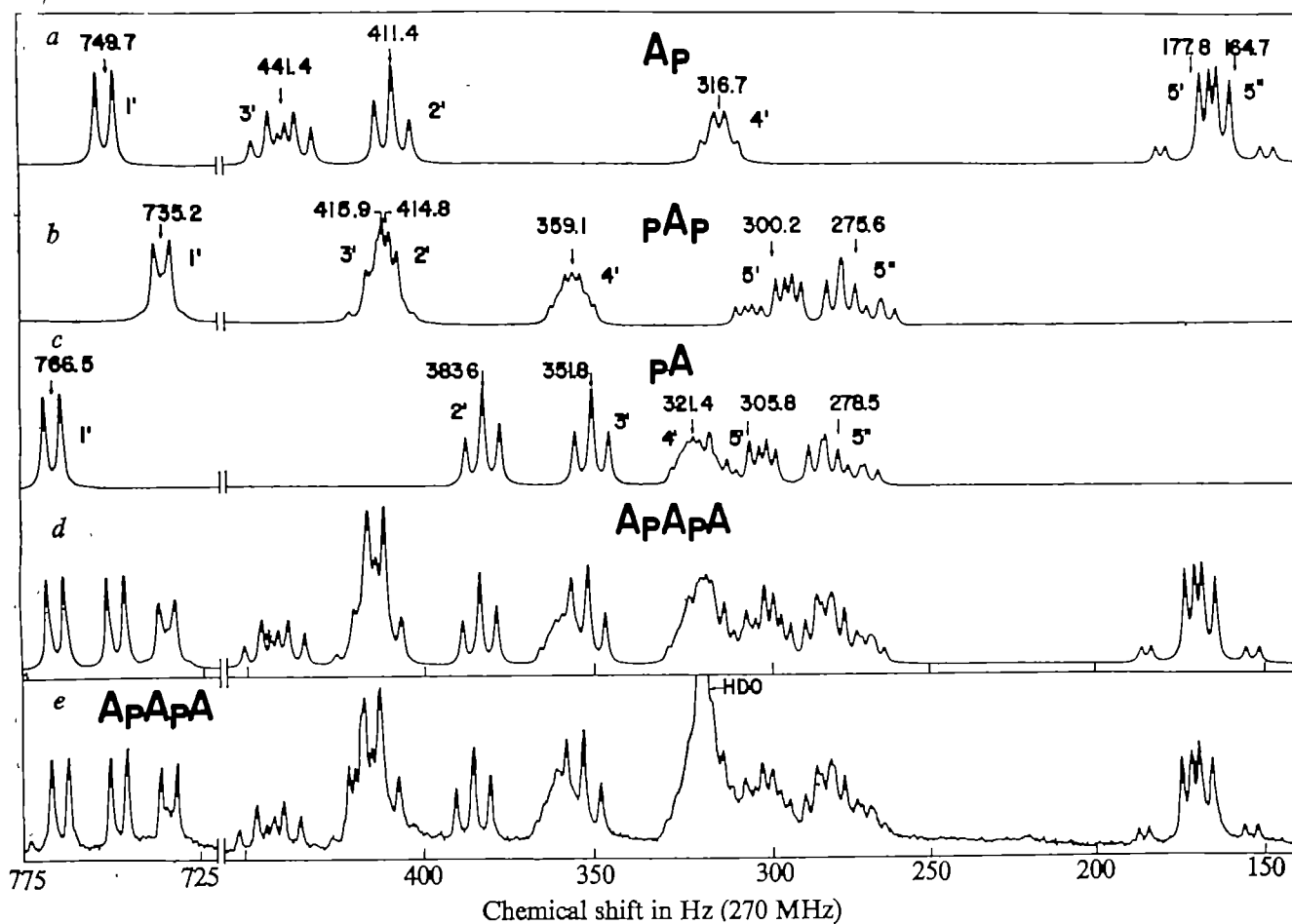
Fig. 2 *a, b, c,* Computer simulations of the Ap–, –pAp– and –pA parts of ApApA. *d,* Combination of the above three parts into one to produce ApApA simulation. *e,* The 270 MHz ¹H experimental NMR spectrum of ApApA at 72 °C, *p*D 7.0, 0.02 M.

values of $J_{H5'-P3'}$ indicate that in the component monomers, as well as the dimer and trimer at 72 °C, the C3'–O3' bonds have similar time-average conformations.

The ribofuranose ring conformation in nucleic acids and their components in aqueous solution may be treated as a simple C2'-*endo* (²E) and C3'-*endo* (³E) equilibrium[22,23] or as an equilibrium of entities with precise phase angles and amplitudes of pucker[24]. The pseudorotational approach[24] is a major improvement over the classical description for solid state data. However, for the majority of nucleotide systems in aqueous solution the pseudorotation approach may not be sufficiently accurate to be an improvement over the simple qualitative approach of a ²E ⇌ ³E equilibrium[23,24]. The percentage population of ³E conformers in adenosine, 3'-AMP, 5'-AMP, 3',5'-ADP, ApA 72 °C, ApApA 72 °C were computed (Fig. 3*a*) and the data show that the component monomers, irrespective of whether a nucleoside or a nucleotide, have the same time-average conformation for the ribose ring. These are only slightly different from that of ApA 72 °C and ApApA 72 °C. Later discussion will show that the ApA and ApApA ribose rings did not approach the same conformational distribution of the components because at the temperature used (72 °C both ApA and ApApA display a certain degree of stacking in D₂O (refs 27, 28).

Two conclusions emerge from the data: (1) In the adenosine series, in aqueous solution the nucleoside and the nucleotides have the same time-average conformation with respect to the ribose ring, the C5'–O5', C4'–C5' and C3'–O3' bonds. It does not seem therefore, that this part of the nucleotide backbone is conformationally more rigid than in nucleosides, at least for the system investigated. The observation that the phosphate group present at the 3' or 5' position or at both the positions

has no detectable effect on the population distribution of conformers about C4'–C5', C5'–O5', C3'–O3' and the ribose ring argues against an effect of the phosphate group on conformation. The mononucleotide components 3'-AMP, 5'-AMP and 3',5'-ADP essentially conserve their backbone conformation as they become integrated into the framework of the destacked oligomers ApA and ApApA.

It is of interest to determine whether these conclusions can be extended to the case of a polymer such as poly(A). The spectrum of poly(A) in the destacking solvent DMSO/D₂O at 86 °C was taken and the entire spectrum was computer simulated (Fig. 4). The extracted values for Σ, Σ' and $J_{H3'-P3'}$ (Table 1) as well as the computed percentage of ³E population (Fig. 3*a*) clearly reveal that destacked poly(A) has a time-average backbone conformation very similar to that of the component monomeric units.

Conformation in the stacked oligomer and polymer

The discussion presented shows that the monomeric units conserve their preferred conformation for the backbone in the destacked oligomer and polymer. This is not true in the case of stacked molecules. Comparison of the computed percentage populations of ³E conformers for stacked poly(A), ApApA and ApA (low temperature in D₂O) with that for the component monomers (Fig. 3*b*) reveals that the ribose ring in the mostly stacked oligomers and polymer displays a conformational preference dramatically different from that of the monomeric components. The ³E populations in the monomers are about 40%, while the corresponding populations for ApA, ApApA and poly(A) in conditions which favour stacking lie in the
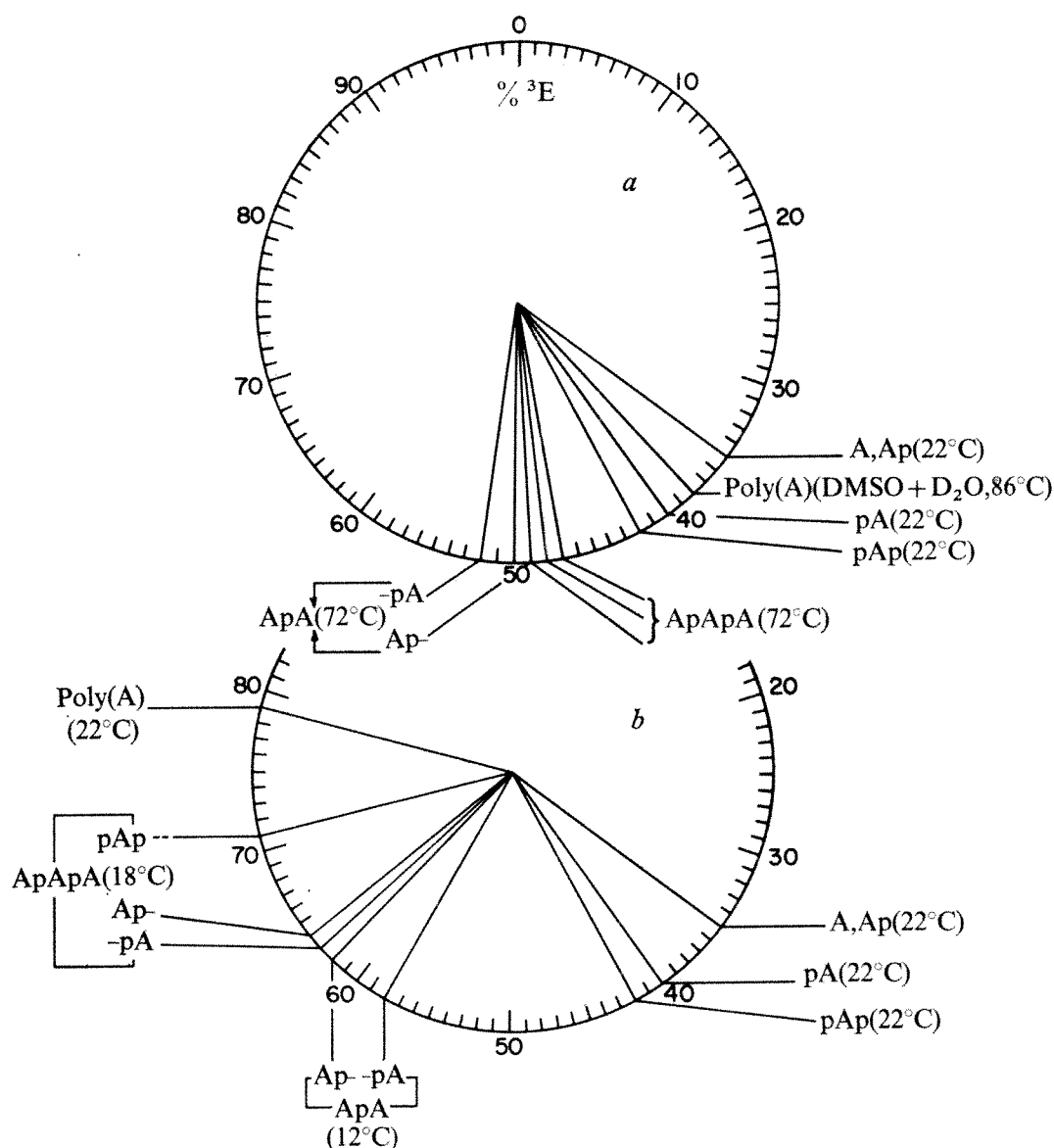
Fig. 3 *a*, Computed percentage populations of the ³E ribose conformers in monomers and in mostly destacked ApA, ApApA and poly(A). *b*, Computed percentage populations of the ³E ribose conformers in monomers and in mostly stacked ApA, ApApA and poly(A). The populations are computed by the method described elsewhere (ref. 15, and F. S. Ezra, C. H. Lee, N.S. Kondo, S. S. Danyluk, and R.H.S.). This method enables us to calculate %³E or %²E populations from either $J_{3'4'}$ or $J_{1'2'}$ using an empirical sum of $J_{1'2'} + J_{3'4'} = 9.5$ Hz. This sum was derived from 30 ribose units of 15 different dinucleoside monophosphates (ref. 15, and F. S. Ezra, C. H. Lee, N. S. Kondo, S. S. Danyluk, and R.H.S.). In the crystals of ApApA, the ribose assumes ³E pucker[26]. The abbreviations A, Ap and pA refer to adenosine, 3'-AMP, 5'-AMP and 3', 5'-ADP. The percentage population of ²E conformer = 100 − %³E.

range of 60–80%. The higher end of the range is preferred by the central –pAp– unit of ApApA and the nucleotidyl units in poly(A). It should be stressed that as the monomeric components become integrated into the framework of a biologically functional stacked polymer, not only is there a significant increase in the populations of ³E conformers, but also there is a shift in the kind of pucker they prefer, that is, the monomers prefer ²E sugar pucker ($\simeq 60\%$) but the oligomers and polymer prefer ³E sugar pucker. Mononucleotides do not conserve their isolated individual conformations as they become part of a polymer in the case of base-stacked adenyl polynucleotides. The low temperature ApApA and poly(A) spectra are extremely complex and so cannot be completely analysed to obtain the values for $J_{H3'-P3'}$, Σ and Σ′. However inspection of the data in Table 1 for 3'-AMP, 5'-AMP, 3', 5'-ADP and ApA (12 °C) shows that stacking significantly changes the conformational preference about the C5'–O5' bond of the –pA unit. On the basis of the temperature dependence of Σ and Σ′ values in ApA, we have previously suggested that the C4'–C5' and C5'–O5' bonds are predominantly *gauche–gauche* in the totally base-stacked state[29].

Conformation in oligo- and polynucleotides other than the adenosine family

Published ribose coupling constant data[12,30,31] on guanosine, 3'-GMP and 5'-GMP in aqueous solution clearly reveal that the nucleosides and the nucleotides display similar distribution of ²E and ³E conformers. The Σ values[12,30,31] show that in all three cases the percentage population of the *gauche–gauche* (ψ = 60°) conformer about C4'–C5' is about 70–75%. Unusual line broadening complications prevent the analyses of the ¹H NMR spectra of GpG and GpGpG and one cannot at present determine the conformational status of the guanosine nucleotidyl units in the oligomer.

Single-stranded polypurine nucleotides are known to exhibit considerably more base stacking than poly(U) in aqueous solution. The coupling constant data for 3'-UMP, 5'-UMP, UpU and poly(U) (Table 2) are, within error limits, similar in their magnitudes, which indicates that in aqueous solution at room temperature the backbone conformation of poly(U) is as flexible as that of the constituent mononucleotides, unlike the case in the adenyl series discussed earlier.

Table 2 Coupling constant data for the various nucleic acid systems of the uridine family

Coupling nuclei	Uridine	3′–UMP	5′–UMP	UpU Up–	UpU –pU	poly(U)*
H1′–H2′	4.4	5.0	4.8	4.2	3.9	5.5
H2′–H3′	5.3	5.2	5.0	5.2	5.1	5.5
H3′–H4′	5.5	5.2	4.3	5.3	5.0	4.0
H3′–P3′	–	8.2	–	8.2	–	9.0
H4′–H5′	3.0	2.8	2.4	2.4	2.4	–
H4′–H5″	4.4	4.2	2.8	4.0	3.2	–
Σ	7.4	7.0	5.2	6.4	5.7	7.5
H4′–P5′	–	–	2.1	–	2.0	–
H5′–P5′	–	–	4.3	–	4.4	–
H5″–P5′	–	–	5.1	–	4.5	–
Σ	–	–	9.4	–	8.9	8.0
H5′–H5″	–12.7	–12.8	–11.7	–13.1	–11.9	–12.5

The concentration is 0.05 M for the monomers. The pD of 3′–UMP and 5′–UMP solutions were kept at pD 5.4 so that the phosphate will be mostly monoanionic like the oligomers and polymer at pD 7.4. For UpU concentration 0.02–0.03 M, pD 7.4. The temperature in all cases was 22 °C. Error in coupling constants ±0.1 Hz unless otherwise stated.
*From ref. 32.

Comparison of the values of $J_{H3'-H4'}$ in uridine and 5′–UMP indicates that this coupling constant undergoes a reduction of 1.2 Hz on phosphorylation at the 5′ position. This translates into an increase of about 15% in the population of ^3E conformers for the ribose ring on 5′ phosphorylation. Equally interesting is the observation that the value of Σ decreases by 2.2 Hz on 5′ phosphorylation of uridine, indicating that the presence of the phosphate group increases the *gauche–gauche* (ψ = 60°) populations about C4′–C5′ by a substantial 20–25%. These observations in the uridine series are very different from those of the adenosine and guanosine series and agree with the projections of a conformationally rigid nucleotide[1,2]. Comparison of the coupling constant data for cytidine, 3′–CMP, 5′–CMP and CpC in Lee *et al.*[15] indicate that in the cytidine series, as in the uridine series, the 5′ nucleotide is conformationally more rigid than the nucleoside; but unlike the uridine system, the mononucleotide conformation is not conserved in the dimer.

Relevance to tRNA conformation

From the foregoing discussion it is clear that the conformational properties of common nucleic acid components, oligo- and

Fig. 4 100 MHz Fourier transformed ^1H-(^{31}P) spectrum of poly(A) in 50% d_6-DMSO and 50% D_2O at 86 °C, 0.05 M (top) and the computer simulation (bottom). The extracted coupling constants from simulation are only accurate ±0.5 Hz because of line broadening complications. It should be pointed out that the accuracy would be considerably less if phosphorus decoupling had not been used.
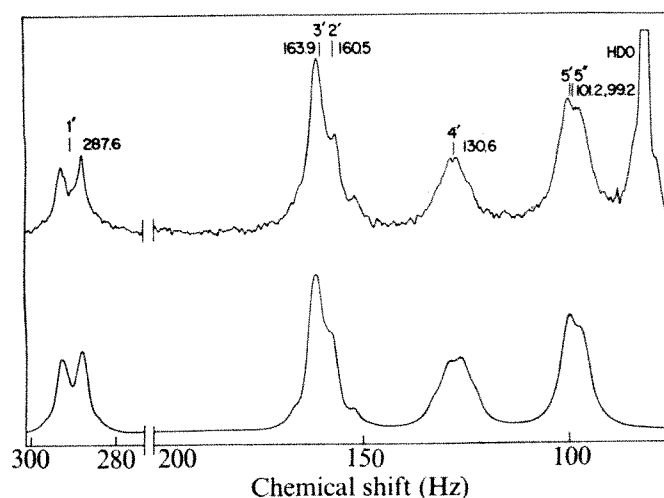
polynucleotides in aqueous solution are not governed by a common principle such as the concept of a conformationally rigid nucleotide[1,2]. On the one hand in the adenosine and guanosine series, the backbones of the nucleosides and nucleotides are equally flexible, but on the other hand in the uridine and cytidine series, the nucleotides are relatively less flexible than their corresponding nucleosides. The concept that the conformational destiny of mononucleotides, as they become part of polynucleotides is ordained by their intrinsic conformations and that they maintain their isolated conformations in the polymer, is true when significant amounts of the polymer are in the destacked state. This is analogous to the solution conformations of coenzyme A and nucleoside diphosphohexoses which exist as a blend of mostly unfolded and folded conformers[14,21]. The conformational conservation breaks down as soon as the mononucleotides become integrated into the backbone framework of a biologically functional base-stacked polynucleotide. The nucleotidyl units in a base-stacked oligomer and polymer exhibit backbone conformations significantly different from the individual conformations of the monomeric components in aqueous solution.

Even though our findings seem to support the reported presence of non-rigid nucleotides in tRNA[33], we wish to examine our results with regard to the sophisticated structure of yeast phenylalanine tRNA derived by Sussman and Kim[34] by examination of the common features from three sets of atomic coordinates. These authors report the presence of extensive base stacking along the central core of the two molecular axes, that is, the amino acid arm and TψC arm make one axis and the dihydrouridine arm and the anticodon arm the other. The bases that are not stacked are D16, D17, G20, U47 and A76. Assuming that the structure of yeast tRNA[Phe] in crystals is the same as that in solution[35–38] from the present findings one may project that in the fully base-stacked regions of tRNA[Phe] in aqueous solutions, the sugar-base torsion will occupy the *anti* domain, the ribose ring a ^3E pucker, the C3′–O3′, C4′–C5′, C5′–O5′, O3′–P and O5′–P preferring domains centred around φ′ ≃ 210°, ψ ≃ 60°, φ ≃ 180°, ω′ ≃ 300° (≃ 60°) and ω ≃ 300° (≃ 60°). The present data do not enable us to make a distinction between right-handed stacks (ω′/ω = 300°/300°) and left-handed loop stacks (ω′/ω = 60°/60°), even though ω′/ω = 60°/60° stacks are expected only in the loop region. One would expect the unstacked region comprising[34] D16, D17, G20, U47 and A76 to show conformational freedom and flexibility so much so that alternative conformers such as ^2E, ψ ≃ 180°/300° become allowed. Because of the demonstrated conformational nexus in aqueous solution among sugar–base torsion, sugar-pucker, C3′–O3′ torsion and ω′/ω rotations (refs 15, 39; F. S. Ezra, C. H. Lee, N. S. Kondo, S. S. Danyluk, and R. H. S., unpublished, and D. M. Cheng and R. H. S., unpublished) one would also anticipate that those fractional populations which exist in ^2E conformations in the unstacked region will also populate in domains in which ω′/ω ≃ 180°/300°, φ′ ≃ 270° and the value of χ, though still in the *anti* domain, several degrees above the ones encountered in the base stacked region. Our general conclusion is that in aqueous solution, conformations of the individual nucleotidyl units of yeast tRNA[Phe] are confined to the classically stable domains, as has been reported in crystals by Sussman and Kim[34] and that non-rigid flexible structures are confined to that region of tRNA which is not base stacked.

¹ Sundaralingam, M., *The Jerusalem Symposia on Quantum Chemistry and Biochemistry*, 5, 417–455 (1973).

2 Sundaralingam, M., in *Structure and Conformation of Nucleic Acids and Protein-Nucleic Acid Interactions* (edit. by Sundaralingam, M., and Rao, S. T.), 487–524 (University Park Press, Baltimore, 1975).
3 Berthod, H., and Pullman, B., *FEBS Lett.*, 30, 231–235 (1973).
4 Young, D. W., Tollin, P., and Wilson, H. R., *Nature*, 248, 513–514 (1974).
5 Viswamitra, M. A., Seshadri, T. P., and Post, M. L., *Nature*, 258, 542–544 (1975).
6 Neidle, S., Kühlbrandt, W., and Achari, A., *Acta Crystallogr.*, B32, 1850–1855 (1976).
7 Reddy, S. B., and Viswamitra, M. A., *Cryst. Struct. Commun.*, 2, 9–13 (1973).
8 Seeman, N. C., Day, R. O., and Rich, A., *Nature*, 253, 324–326 (1975).
9 Tsai, C., Jain, S. C., and Sobell, H., *Proc. natn. Acad. Sci. U.S.A.*, 72, 628–632 (1975).
10 Evans, F. E., and Sarma, R. H., *Biopolymers*, 13, 2117–2132 (1974).
11 Evans, F. E., and Sarma, R. H., *Cancer Res.*, 35, 1458–1463 (1975).
12 Evans, F. E., and Sarma, R. H., *J. Am. chem. Soc.*, 97, 3215–3218 (1975).
13 Sarma, R. H., and Mynott, R. J., *J. Am. chem. Soc.*, 95, 1641–1649 (1973).
14 Lee, C. H., and Sarma, R. H., *Biochemistry*, 15, 697–704 (1976).
15 Lee, C. H., Ezra, F. S., Kondo, N. S., Sarma, R. H., and Danyluk, S. S., *Biochemistry*, 15, 3627–3639 (1976).
16 Kondo, N. S., Ezra, F., and Danyluk, S. S., *FEBS Lett.*, 53, 213 (1975).
17 Borer, P. N., Karr, L. S., and Ts'o, P. O. P., *Biochemistry*, 14, 4847–4863 (1975).
18 Sarma, R. H., Lee, C. H., Evans, F. E., Yathindra, J., and Sundaralingam, M., *J. Am. chem. Soc.*, 96, 7337–7348 (1974).
19 Lee, C. H., and Sarma, R. H., *J. Am. chem. Soc.*, 98, 3541–3548 (1976).
20 Sarma, R. H., Mynott, R. J., Wood, D. J., and Hruska, F. E., *J. Am. chem. Soc.*, 95, 6457–6459 (1973).

21 Lee, C. H., and Sarma, R. H., *J. Am. chem. Soc.*, 97, 1225–1236 (1975).
22 Sarma, R. H., and Mynott, R. J., *J. chem. Soc.*, Chem. Commun., 975–976 (1972).
23 Evans, F. E., and Sarma, R. H., *J. biol. Chem.*, 249, 4754–4759 (1974).
24 Altona, C., and Sundaralingam, M., *J. Am. chem. Soc.*, 95, 2333–2344 (1973).
25 Davies, D. B., and Danyluk, S. S., *Biochemistry*, 13, 4417–4434 (1974).
26 Suck, D., Manor, P. C., Germain, G., Schwalke, C. H., Weimann, G., and Saenger, W., *Nature new Biol.*, 246, 161–165 (1973).
27 Chan, S. I., and Nelson, J. H., *J. Am. chem. Soc.*, 91, 168–178 (1969).
28 Ts'o, P. O. P., Kondo, N. S., Schweizer, M. P., and Hollis, D. P., *Biochemistry*, 8, 997–1029 (1969).
29 Evans, F. E., Lee, C. H., and Sarma, R. H., *Biochem. biophys. Res. Commun.*, 63, 196–214 (1975).
30 Davies, D. B., and Danyluk, S. S., *Biochemistry*, 14, 543–554 (1975).
31 Lee, C. H., Evans, F. E., and Sarma, R. H., *J. biol. Chem.*, 250, 1290–1296 (1975).
32 Chan, S. I., and Kreishman, G. P., *J. Am. chem. Soc.*, 92, 1102–1103 (1970).
33 Jack, A., Klug, A., and Ladner, J. E., *Nature*, 261, 250–251 (1976).
34 Sussman, J. L., and Kim, S. H., *Science*, 192, 853–858 (1976).
35 Kim, S. H., *Prog. Nucleic Acid Res. molec. Biol.*, 17, 181 (1976).
36 Sigler, P. B., *Am. Rev. Biophys. Bioengng*, 4, 477 (1975).
37 Langlois, R., Kim, S. H., and Cantor, C. R., *Biochemistry*, 111, 2554 (1975).
38 Robillard, G. T., Tarr, C. E., Vosman, F., and Berendsen, H. J. C., *Nature*, 262, 363–369 (1976).
39 Singh, H., Herbert, M. H., Lee, C. H., and Sarma, R. H., *Biopolymers* (in the press).

letters to nature

Detection of hot gas in clusters of galaxies by observation of the microwave background radiation

THE Uhuru (ref. 1) and the Ariel V (ref. 2) satellites have shown that many rich clusters are powerful sources of X rays. This emission has been interpreted either as thermal bremsstrahlung of very hot gas filling the clusters[3,4], or as inverse Compton scattering of photons by relativistic electrons[5,6]. Spectral evidence[2] has begun to favour a thermal origin for this radiation, implying the existence of large amounts of hot gas. Such gas would have a profound influence upon the dynamics of radio sources, and its mass might well be a substantial fraction of that of the whole cluster. Hot gas may therefore be a major constituent of the Universe, so that independent confirmation of its existence is extremely important. We present here observations of small diminutions in the cosmic microwave background radiation in the directions of several rich clusters of galaxies. This confirms the existence of large amounts of very hot gas in these clusters and indicates that their X radiation is thermal bremsstrahlung and not inverse Compton emission.

A necesary consequence of the existence of hot gas in clusters is the Compton scattering of the cosmic microwave background by hot electrons[7]. As a result, some of the microwave photons are redistributed with slightly higher energies, and the effect may be observed at radio frequencies as a small diminution in the background temperature. The magnitude of this effect is predicted to be $\sim 0.5 \times 10^{-3}$ K for the gas clouds associated with cluster X-ray sources ($\rho_{gas} \simeq 5 \times 10^2$ atom m^{-3}, $T_{gas} \simeq 2.5 \times 10^8$ K, $l \simeq 300$ kpc (refs 4, 7)). Parijskij[8] has claimed to have detected this effect at a slightly higher (1.2×10^{-3} K) level in the direction of the Coma cluster, but his result was marginal and in any case seems to have been implicitly withdrawn[9].

We present the first results of an attempt to detect this effect, using the 25-m telescope at the Chilbolton Observatory of the SRC Appleton Laboratory. We have observed 7 clusters, of which 6 are known to be X-ray sources. The remaining cluster (A2218) is very distant, so that no strong limits to its X-ray flux are available. It was included because it is a very rich cluster (richness 4 on Abell's scale) and was known not to contain any non-thermal radio sources likely to mask a diminution in the intensity of the microwave background.

The Coma cluster (A1656) was included, although it is confused by the weak radio source 5C4.85 (36 ± 6 mJy at 10.6 GHz), for which a small correction has been made. Upper limits on the radio flux density from the other clusters are < 100 mJy at 1,400 MHz, though A478 is only 43' distant from the bright radio source 3C109 and no limit on its flux density has been published.

To minimise the effect of non-thermal background sources, it is desirable to observe at the highest possible frequency compatible with atmospheric stability. A frequency of 10.6 GHz was chosen, at which the Chilbolton aerial has a half-power beamwidth of 4.5'. Observations were made using a twin-beam system, the beams being separated on the sky by 14' in the azimuthal direction, which is just sufficient to take the reference beam beyond the central regions of the largest cluster (Coma). Ideally, the beams should be more widely separated than this, but atmospheric irregularities and distortion of the off-axis beam then become severe problems. The receiver included an uncooled one-stage parametric amplifier, having a noise temperature of ~ 280 K and a bandwidth of ~ 200 MHz. The output time constant was 1 s and samples were taken at 0.5-s intervals. The two beams were alternately directed at the centre of a candidate cluster for 20 s at a time; after a complete cycle the mean (x_i) and standard error (σ_i) of the 80 samples (~ 20 independent samples) were stored on-line in a Ferranti ARGUS 500 computer. This 'wagging' procedure was continued until, at the end of an allotted observation period (~ 4 h), the cumulative mean and s.d. were derived, weighting the individual observations by $1/(\sigma_i^2 + x_i^2/2)$ This latter term proved to be an effective gate against spurious high-significance measurements, caused by either atmospheric irregularities with time scales ~ 1 min, receiver drift, or occasional interference from radar. The system was calibrated at regular intervals by injecting a known noise signal into the main beam. The antenna temperatures measured in this way were then converted into equivalent sky temperatures, using radio sources of known flux density and the measured beamshape to determine the aerial efficiency (55%). We have also attempted observations using drift scans. This method is, however, particularly sensitive to atmospheric irregularities and was abandoned in favour of the wagging technique.

The results are given in Table 1. They represent a total of 674 h of observations in 12 weeks during 1975 Nov–1976 June. They show the temperature of background radiation in the directions of the centres[10] of the 7 Abell clusters, relative to a

Table 1 Data obtained

Cluster	Temperature ($\pm$ s.d.) (10^{-3} K)	$\sqrt{\chi}(^2-\sqrt{})n-1$
A376	-0.13 ± 0.66	1.75†
A478	$+0.33\pm0.52$	-1.09
A576	-0.71 ± 0.57	0.34
Coma*	-1.51 ± 0.40	1.89†
A2218	-1.94 ± 0.54	1.92†
A2319	-0.13 ± 0.41	0.38
A2666	-0.27 ± 0.35	0.42
Blank Sky	-0.01 ± 0.32	-0.39

*Combination of two different positions (see text).
†Slightly discordant.

peculiarly weighted mean of the temperature over arcs 14' on each side of the centre. Each value is the result of many observation periods, and statistical tests have shown that the data are generally concordant and conform rather well to the expected Gaussian distribution. All results have $\sqrt{\chi^2}-\sqrt{(n-1)}<2$, where n is the number of observation periods, so that considerable reliance can be placed on the final standard errors. The result for the Coma cluster is the weighted mean of two different positions, one of which is coincident with the centre of the cluster ($\alpha = 12$ h 57 min 24 s, $\delta = 28°15'$) and is only 3.16' from 5C4.85. A correction of $-1.35\pm0.25 \times 10^{-3}$ K has been calculated from the measured beamshape and flux density of 5C4.85 and applied to this result. No significant correction is required for the second position ($\alpha = 12$ h 57 min 37 s, $\delta = 29°13'$). For comparison, the uncorrected average value for coma is $-1.02 \pm 0.40 \times 10^{-3}$ K. The 'blank sky' observation was included to guard against the possible existence of systematic errors. The value shown is the weighted mean of three different positions, each chosen to have the same declination as either A576 or A2218, thereby ensuring that the observing conditions were nearly identical to those prevailing for the clusters themselves. No evidence of systematic errors was found. At our observing frequency, confusion from non-thermal background sources is expected to be $\sim 0.1 \times 10^{-3}$ K (J. V. Wall, personal communication).

It will be seen that, apart from A478, the results are in the expected sense and amount to a marginal detection of the expected diminutions at the predicted level of 0.5×10^{-3} K. The formal probability of obtaining results of this significance merely by chance is low, as may be seen from the value of χ^2 ($\chi^2 = 30.0$) for all 7 clusters assuming that the effect is absent. With 7 degrees of freedom, this indicates that the results are inconsistent with the assumption of zero mean at the 99.99% confidence level. This test takes no account, however, of the relevant fact that 6 out of 7 results are in the same sense. A further test is the value of the weighted mean of all the clusters, which is $-0.63\pm0.17\times10^{-3}$ K. It must be realised that this figure is only relevant as a test of the null hypothesis and has no particlur astrophysical significance. It should also be emphasised that no data have been rejected and that Table 1 is a complete list of clusters for which we have results. Bearing this in mind, it is clear that the results are most encouraging, if not yet entirely conclusive. Observations with greater sensitivity are now proceeding, in an attempt to confirm the effect at a higher confidence level, and to measure the shape of the diminution for one cluster.

These observations therefore strongly suggest the presence of hot plasma ($> 10^8$ K) filling clusters of galaxies. It seems highly prob·' 'e that the X radiation from these clusters is just the therm.' ·· ‑ahlung of this gas, and that previous inverse Compton interpretations can now be ruled out. Confirmation of these effects will also show directly that the microwave background radiation is truly cosmic, and must have its origin beyond distant clusters of galaxies.

S. F. GULL

K. J. E. NORTHOVER

Mullard Radio Astronomy Observatory,
Cavendish Laboratory,
Madingley Road, Cambridge CB3 0HE, UK

1 Giaconni, R., *et al.*, *Astrophys. J. Suppl.*, 27, 37 (1974).
2 Mitchell, R. J., Culhane, J. L., Davison, P. J. N., and Ives, J. C., *Mon. Not. R. astr. Soc.*, 175, 29P (1976).
3 Lea, S. M., Silk, J., Kellogg, E., and Murray, S., *Astrophys. J. Lett.*, 184, L105 (1973).
4 Gull, S. F., and Northover, K. J. E., *Mon. Not. R. astr. Soc.*, 173, 585 (1975).
5 Brecher, K., and Burbidge, G. R., *Nature*, 237, 440 (1972).
6 Okoye, S. E., *Mon. Not. R. astr. Soc.*, 160, 339 (1972).
7 Sunyaev, R. A., and Zel'dovich, Ya.B., *Comments Astrophys. Space Phys.*, 4, 173 (1972).
8 Parijskij, Y. N., *Astr. Zh.*, 49, 1322 (1973).
9 Parijskij, Y. N., *Astrophys. J. Lett.*, 180, L47 (1973); (see also correction 188, L122; 1974).
10 Sastry, G. N., and Rood, H. J., *Astrophys. J. Suppl.*, 23, 371 (1971).

Visibility of galaxies

IT is well known that our counts of galaxies could be seriously biased by selection effects, largely determined by the brightness of the night sky. To illustrate this, suppose the Earth were situated near the centre of a giant elliptical galaxy, then the mean surface brightness of the sky would appear some 8–9 mag brighter than is observed from our position in the Galaxy (~ 23 V mag (arc s)$^{-2}$ looking toward the galactic pole, discounting atmospheric and zodiacal contributions[1,2]). Optical astronomers would then find extragalactic space an empty void; spiral and irregular galaxies would be quite invisible and all they would easily detect of galaxies would be the core regions of ellipticals very similar to their own. They would be blinded to much of the Universe by the surface brightness of their parent
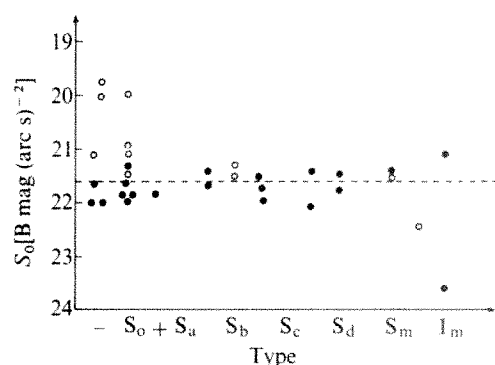
Fig. 1 Central surface brightnesses S_0 of exponential disks of spiral and irregular galaxies taken from Freeman[2] plotted against galactic type. Filled circles show most reliable data, open circles less reliable. Some of the original points, subsequently shown to be incorrect (K. C. Freeman, personal communication) have been removed. (— — — —) is at 21.65 B mag (arc s)$^{-2}$.

galaxy. But this blinding is clearly a relative matter and we should ask to what extent we are blinded by the spiral galaxy in which we exist, faint as it may appear by comparison. I will argue that strong indirect evidence already exists that our knowledge of galaxies is heavily biased by the sky background, and that the true population of extra-galactic space may be very different from the one we can see.

Begin by recalling two sets of observations applying to spirals and ellipticals respectively, that have been in the literature some time. Both are based on de Vaucouleurs' discovery[3] that the radial surface brightness distribution $\sigma(r)$ of galaxies (apart from second order effects like spiral arms) can be described by the law

$$\log_{10}\left[\frac{\sigma(r)}{\sigma(0)}\right] = -\left(\frac{r}{\alpha}\right)^{1/\beta} \qquad (1)$$

where α is a scale-length and $\beta = 1$ for spirals and 4 for ellipticals. Using equation (1) one can show that the total luminosity L_T is

$$L_T = L(r\rightarrow\infty) = \int_0^\infty 2\pi r \sigma(r)\,dr = \frac{(2\beta)!}{(\log_e 10)^{2\beta}}\,\pi\sigma(0)\alpha^2 \qquad (2)$$

In 1970 Freeman[2] collated the best photometry then available for spiral and irregular systems and found the surprising result that $\sigma(0)$ was virtually constant for the disks of all but a few
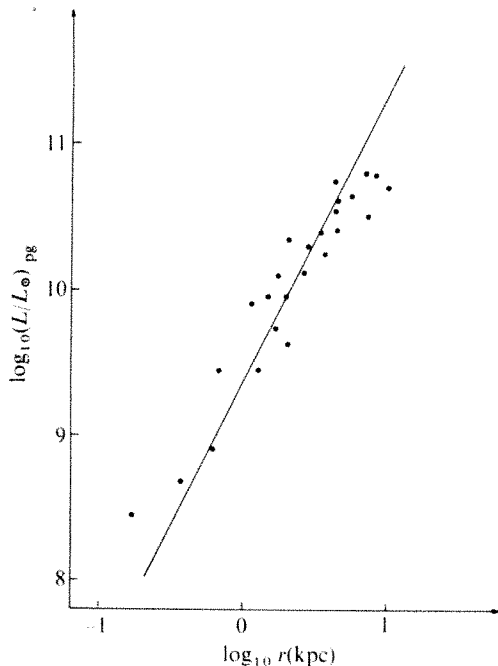
galaxies in the sample. If we express $\sigma(0)$ in B mag (arc s)$^{-2}$ and then write it as S_0, then Freeman finds $\langle S_0\rangle_s = 21.65 \pm 0.3$ (see Fig. 1).

Fish[4] had earlier discovered an apparently unrelated fact about a sample of well studied ellipticals: he showed that their binding energy Ω varied as $M^{3/2}$ where M is their mass. Fish actually measured L and α for his sample, and then assumed that ellipticals generally have the same mass to light ratio. Disentangling the directly measured quantities from the assumptions

$$\frac{\Omega}{M^{3/2}} = \text{const} \rightarrow \frac{GM^2}{RM^{3/2}} = \text{const} \rightarrow \frac{L_T^{1/2}}{R} = \text{const}$$

But R (the radius) $\sim \alpha$ so, using equation (2), Fish's law implies that

$$\frac{L_T}{\alpha^2} \sim \sigma(0) = \text{const}$$

for the ellipticals in his sample, which comprised most of the best studied individuals at that time. Fish's raw data, not plotted in his paper, are shown in Fig. 2, and we see that Fish's law is conceptually identical with Freeman's, and can be stated thus: well studied bright ellipticals have a central surface brightness $\langle S_0\rangle_E = 14.80 \pm 0.9$ B mag (arc s)$^{-2}$.

When the two laws are placed contiguously in this fashion one begins to suspect that observational selection is responsible, though one is puzzled by the large discrepancy in $\langle S_0\rangle_s - \langle S_0\rangle_E = 6.85$ mag between the two galactic types. I now argue that observational selection can indeed account for both Fish's and Freeman's results rather simply, and even supplies the two different values of $\langle S_0\rangle$ with few assumptions beyond equation (1), which is well tested.

Fish's and Freeman's sample comprise the major proportion of galaxies for which careful surface photometry existed at that time. What criteria are used in selecting galaxies for this type of investigation? De Vaucouleurs[5] has discussed this matter but there is no clearcut answer and, historically, apparent size, apparent area, apparent luminosity and surface brightness have probably all been important. The selection has, however usually, been based initially on photographic material. Because of the limited dynamic range of emulsions the real luminosity is not well measured, and to a first approximation at least, galaxies with large angular radii or areas will generally appear most luminous on a photographic plate. (The belated discovery of Zwicky's compacts, many of which are highly luminous, testifies to this natural selection effect.) Our hypothesis is that it is the apparent radius r_{ap} (and area r_{ap}^2) which gives a galaxy a spectacular appearance on a plate, and makes us believe that it is probably luminous and suitable for more detailed investigation. On this hypothesis the galaxies in the Fish–Freeman samples were therefore probably chosen, consciously or otherwise, mainly for their apparent radii and areas in photographic surveys.

For constant L_T and S_L (where S_L is the limiting surface brightness out to which diffuse objects are readily apparent on a photographic plate) intuition leads one to expect that r_{ap} will have a maximum as $\sigma(0)$ varies. If $\sigma(0)$ is low, even the central isophotes will lie close to σ_L and r_{ap} will be small; conversely, if $\sigma(0)$ is high, then by equation (2), α and hence r_{ap} will also be small.

If σ_L is the limiting surface brightness corresponding to S_L (where S_L is σ_L expressed in magnitudes) then equation (1) becomes

$$\log_{10}\left[\frac{\sigma_L}{\sigma(0)}\right] = -\left(\frac{r_{ap}}{\alpha}\right)^{1/\beta}$$

Then:

$$r_{ap} = \alpha(0.4)^\beta (S_L - S_0)^\beta \qquad (3)$$

Using equation (2) to eliminate α from (3) we obtain

$$r_{ap} = KL_T^{1/2}\left[\frac{1}{\pi 2\beta!}\right]^{1/2} (0.4\log_e 10)^\beta \exp(0.46\,S_0)(S_L - S_0)^\beta \qquad (4)$$

where K is a normalising constant whose value can be obtained from any well studied galaxy, elliptical or spiral. Using NGC3379 (refs 5, 6) with r_{ap} in kpc and L_T in units of $10^9 L_\odot$, $K = 1.20 \times 10^{-4}$ and the results are plotted in Fig. 3. As anticipated there are well defined maxima in r_{ap} for the two types of galaxy. What is surprising is that the $S_0(r_{ap} = \text{max})$ are separated for the spirals and ellipticals by 6.52 mag, which is in the right direction and

very close to the precise size required to explain both Fish and Freeman's laws. Moreover, since for most plate material of interest here, S_L must be close to[7] 24 B mag (arc s)$^{-2}$, then $S_0(r_{ap} = max) = 21.83$ for spirals and 15.31 for ellipticals, so the absolute values are close to those observed as well. Furthermore the scatter in the spiral $S_0(r_{ap} = max)$ would be smaller than for ellipticals as the observations demand. There are no free parameters in Fig. 3 for the β follow from equation (1) and L_T changes only the vertical scale (see equation (4)). It is important to note that L_{ap}, where L_{ap} is the apparent photographic luminosity, also has a double peaked distribution similar to r_{ap} in Fig. 3, but the details are complex and for brevity we omit discussion of L_{ap} here.

What is one to make of the uncanny fit of Fish's and Freeman's data to equation (4)? Conservatively, one could argue that their data are not representative and that the fit is an unlikely coincidence, but a coincidence none the less. Though this could well be correct it is not an heuristically useful hypothesis and we pursue it no further here.

The alternative is to accept our hypothesis, and acknowledge that surprisingly powerful selection effects are involved in picking out galaxies which are apparently large and apparently luminous, and that only objects with a very narrow range of σ(0) will be so selected. If this is true, and only further observations can prove it to be so, then it is likely that our present knowledge of the galactic luminosity function is seriously in error, in that galaxies of all luminosities with σ(0) different from
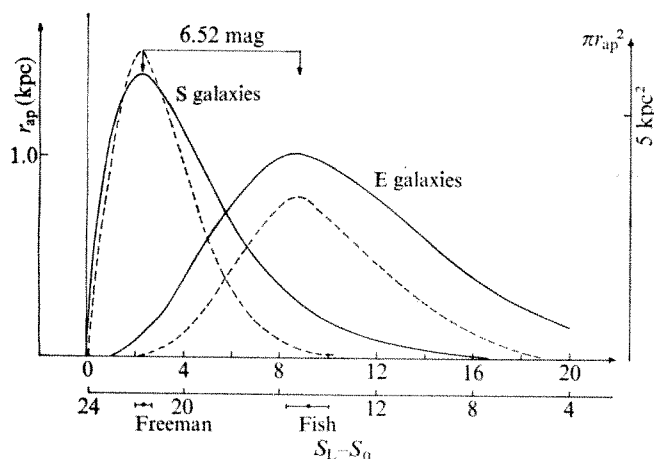
Fig. 3 Apparent radii r_{ap} (solid lines) and areas πr_{ap}^2 (dashed) for ellipticals and spiral disks of total luminosity $L_{pg} = 10^9 L_{\odot pg}$, as a function of $(S_L - S_0)$. Radius scale on left, area on right. Changing L varies only the vertical scale. The scale under the abscissa is based on the assumption that $S_L = 24$ B mag (arc s)$^{-2}$.

the Fish–Freeman values will be under-represented. Stated thus the hypothesis is hardly new. At one end of the scale Zwicky's work has revealed the hitherto unsuspected compacts while Arp[8], for instance, has demonstrated that the log L/log r plane is everywhere populated where it is detectable, and that there are no good reasons for believing that lower surface brightness objects, and parts of objects, do not exist. Fish's and Freeman's laws are then strong, but indirect evidence for this hypothesis.

On this hypothesis galaxies are like icebergs and what is seen above the sky background may be no reliable measure of what lies underneath. An apparently insignificant dwarf may be the core of a large galactic system of either type. The resolved Sculptor/Fornax systems have S_0 as low as 25 B mag (arc s)$^{-2}$ so the whole range of S_0 between 25 and 15 could be populated. Conceivably the missing mass in groups, clusters and the Universe may be accounted for in this way.

To test the selection hypothesis it will be necessary to move away from the narrowly selected range of objects with spectacu-

lar apparent radii. In a given region of sky all objects suspected to be non-stellar will have to be studied in detail. Since galaxies often contain both elliptical and exponential components it will be necessary to measure both S_0 and α for each component before L_T can be established.

Such a systematic attack could take years. In the shorter term it is interesting to wonder if some few icebergs may be turned up by more *ad hoc* approaches:

(a) Figure 3 reveals that the most dramatic gradient in r_{ap} occurs for spirals with small $S_L - S_0$. The S.R.C. IIIaJ survey should have a deeper S_L than the IIaO and in some cases faint spirals on the IIaO should show a dramatic increase in r_{ap} in going to the IIIaJ.

(b) The cores of both spirals and ellipticals commonly obey the de Vaucouleurs elliptical law. Insignificant-looking ellipticals may sometimes be the tips of giant low-σ(0) spirals which should therefore contain hydrogen. A quick 21-cm survey of a large number of dE might be interesting since Fisher and Tulley's sample[9] contains few dE (dwarf ellipticals).

(c) Although the spiral and H II regions of a galaxy generally represent only a small fraction of the total light, they do have a higher contrast against the night sky. Deep Schmidt plates at wavelengths designed to enhance this contrast (for example Hα) may show some spiral arms with no underlying disk. Ring galaxies may be objects of this type.

(d) Some systems show signs of tidal interference with no apparent cause. 'Dwarfs' in the vicinity warrant careful examination.

(e) Galaxies are often a composite of the two types. Take a disk with $S_0 = 21.65$ mag and add to it an elliptical component with the same S_0 and L_T; equation (4) gives the elliptical r_{ap} as 7% of the disk r_{ap} and so the core will appear insignificant, though it represents the tip of a massive halo. It may be possible to detect such halos by careful measurement of the core scale-length.

(f) Double radio sources with 'empty fields' may sometimes have at their centre a giant galaxy of low S_0, which may be detectable with careful photometry.

In opposition to the selection hypothesis in its most radical form we should point out that σ(0) is not likely to be independent of the total mass, M, of the galaxy, and the giants (that is high mass objects) will in general have a higher σ(0) and will not be so easily missed. But if we suppose that M and $\bar{\rho}$ are the important physical parameters for classifying galaxies then on dimensional grounds σ(0) ~ $M^{1/3}\bar{\rho}^{2/3}$. σ(0) for a giant of $10^{11}M_\odot$ will then in general be less than 3 magnitudes higher than for an extreme dwarf of $10^8 M_\odot$ which is a small difference compared to the total observed range of S_0 between ~8 for extreme compacts[10] and 25 for Sculptor–Fornax systems, so this is probably not an important effect.

Observations are urgently required to determine how important surface brightness selection may be.

I am most grateful to Ken Freeman for many stimulating and instructive conversations.

M. J. DISNEY

Kapteyn Astronomical Institute,
University of Groningen,
The Netherlands

1 Sears, F. H., *Astrophys. J.*, 52 162 (1920).
2 Freeman, K. C., *Astrophys J.*, 160, 811 (1970).
3 de Vaucouleurs, G., *Handbuch der Physik*, 53, 311 (1959).
4 Fish, R. A., *Astrophys. J.*, 139, 284 (1964).
5 de Vaucouleurs, G., *IAU Symp.*, 58, 1 (1974).
6 Miller, R. H., and Prendergast, K. H., *Astrophys. J.*, 136, 713 (1962).
7 Heidmann, J., Heidmann, N., and de Vaucouleurs, G., *Mem. R. astr. Soc.*, 75, 85 (1971).
8 Arp, H., *Astrophys. J.*, 142, 402 (1965).
9 Fisher, J. R., and Tulley, R. B., *Astr. Astrophys.*, 44, 151 (1975).
10 Sargent, W. L. W., *Astrophys. J.*, 160, 405 (1970).

Trends in the climate of the North Atlantic Ocean over the past century

THE dominant large scale phenomena determining the climate of the North Atlantic Ocean are the current systems of the Gulf Stream and the North Atlantic Drift as well as the pattern of atmospheric circulation associated with the semi-permanent centres, the Iceland low and the Azores high. Recently, three long time series of data have become available which throw some light on the interactions between these ocean–atmosphere systems and which appear to indicate the importance of advective processes, as opposed to upwelling, in determining long term variations in the surface temperature of the ocean.

Interactions between the ocean currents and the atmospheric circulation have been studied by Brooks[1] and subsequently by Bjerknes[2]. The latter postulated that long-term fluctuations in the sea-surface temperature of the North Atlantic were determined by such interactions and, in particular, that periods of low sea-surface temperature corresponded with periods of intensification of the Iceland low and the Azores high and, hence, of the mid-latitude westerly winds. The increased wind was assumed to produce an intensification of the flow of the Gulf Stream. Bjerknes considered that a drop in surface temperature of the northern North Atlantic may be attributed to upwelling

produced by increased cyclonic wind-drag and also to an increase in the cold winds from North America. Working with data from the North Atlantic weather stations from 1952 to 1965, Rodewald[3] has confirmed that cooling of the surface waters of the ocean is associated with an intensification of the atmospheric circulation.

The three long-time series of data referred to above are :

(1) Data for sea-surface temperature, based largely on observations from ships in transit, are held in the marine data deck of the US National Climatic Centre. Monthly means of sea-surface temperature for each 5° × 5° quarter of Marsden squares 145 and 182 (see Fig. 1) have been made available to the author by the Pacific Environmental Group of the US National Marine Fisheries Service. These areas of the western approaches to Great Britain were selected because they have been well covered by ships and virtually complete monthly data are available from 1854 to 1968, updated to 1973 by data supplied by the UK Meteorological Office.

(2) As an index of changes in the strength of flow of the Gulf Stream, Martin[4] has recently presented a time series of data for mean sea level at points along the coast of Florida, for the period 1902–64.

(3) Cohen and Sweetser[5] have provided a time series of freqencies of tropical cyclones in the Atlantic for 1871–1973. The positions and intensities of the semi-permanent centres reflect the conditions for the formation and movement of hurricanes and Namias[6] has shown that a southward
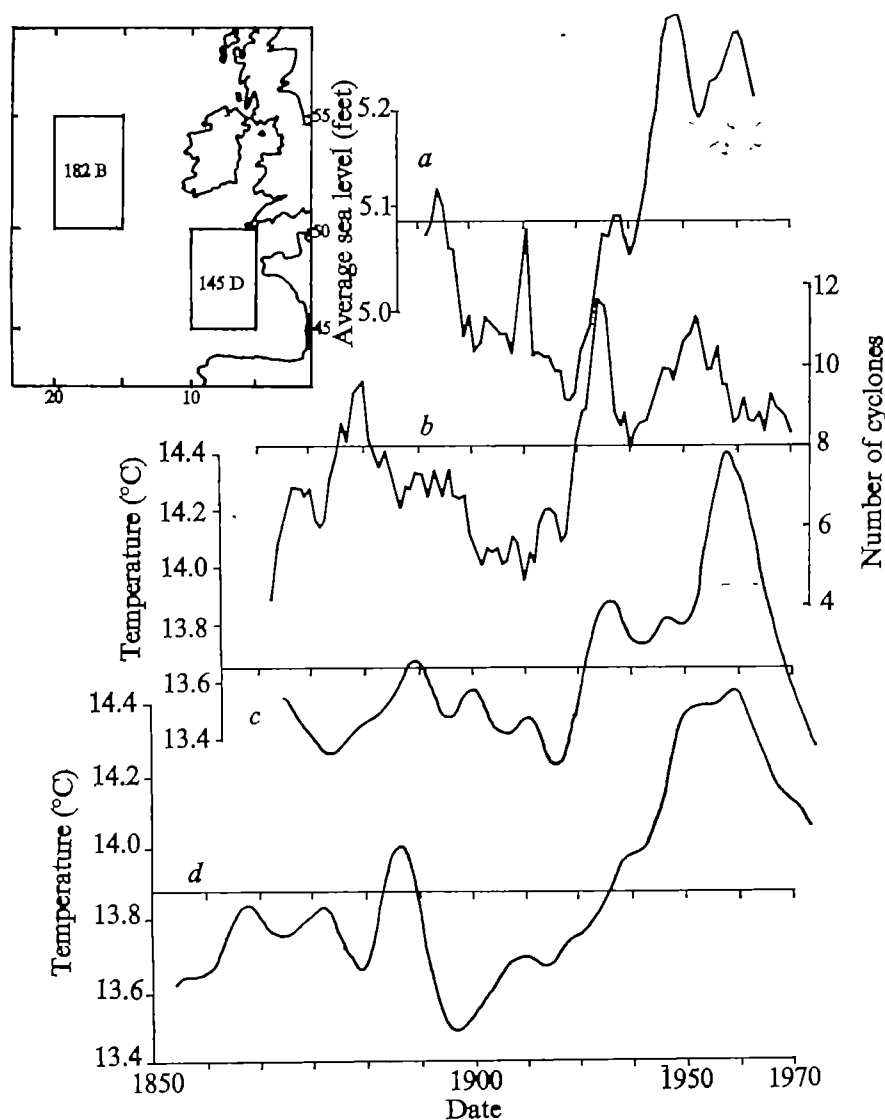
Fig. 1 *a*, 5-yr running mean of average sea level at stations along the coast of Florida, redrawn from Martin[4]. *b*, 7-yr running mean number of Atlantic tropical cyclones, redrawn from Cohen and Sweetser[5]. *c*, Smoothed variations of sea-surface temperature for Marsden square 182B (50–55°N, 15–20°W). *d*, Smoothed variations of sea-surface temperature for Marsden square 145D (45–50°N, 5–10°W). The variables in *c* and *d* were both smoothed by the application of eigenvector filters which, in both instances, were approximately equivalent to 7-yr running means. Inset shows position of Marsden squares 145 and 182.

shift of the Azores high is associated with diminished hurricane activity. It follows that the frequency of tropical cyclones can be used as a guide to the long term changes in the intensity of atmospheric circulation over the North Atlantic.

These three time series are illustrated in Fig. 1, all expressed as smoothed variables (see legend). The general similarities between the long term patterns of change are obvious, showing that low sea-surface temperatures in the North Atlantic are associated with a higher-than-normal flow in the Gulf Stream and a lower-than-normal frequency of tropical cyclones, indicating an above average intensity of the mid-latitude westerlies. Martin illustrated the relationship between Gulf Stream flow and sea-surface temperature in an area off the west coast of Scotland (data from Smed[7]) and he considered that this supported the proposal by Iselin[8] that strong flow in the Gulf Stream leads to a retention of warm water in the North Atlantic Gyre and a reduction in the flow of this water in the North Atlantic Drift, with a consequent drop in temperature. This conflicts with Bjerknes' interpretation that cooling is caused by upwelling induced by increased cyclonic wind-drag:

If the Martin–Iselin interpretation, based on advected changes, is correct, then one would expect the cooling and warming trends, measured at the sea surface, to be more pronounced during the winter and spring in the absence of vertical temperature stratification, while cooling by upwelling should be equally or even more pronounced during periods of summer stratification.

Examination of the time series of sea-surface temperature data for individual months in Marsden squares 182 and 145 indicates that, while a common long-term pattern of change is present in each month of the year, it was clearest in the winter and spring months, whereas during stratification (Schroeder[9]) it tends to be obscured by higher frequency variations, presumably because of local fluctuations in radiation. Thus, these data clearly support the hypothesis that temperature change is produced by advective processes as opposed to upwelling.

It would appear that the time series illustrated in Fig. 1 provides a description, in relative terms, of the major climatic elements of the North Atlantic Ocean covering the past hundred years. It must be stressed that these series contain little or no evidence on which to base predictions; moreover, there are shorter-term phenomena of equal or even greater magnitude which influence both the atmospheric circulation and sea temperatures. Nevertheless, it is of interest to compare the long term history of variations in sea-surface temperature with the more recent changes. Wahl and Bryson[10], on the basis of Rodewald's studies of data from the ocean weather stations, state that "it seems that we have experienced, in the past 29 yr or so, a change in Atlantic Ocean temperatures which amount, in the Gulf Stream vicinity, to about one-sixth of the difference between total glaciation and our present climate"; they also mention two other recent changes with equivalent magnitudes: the northward penetration of the monsoon rains in West Africa and a global increase in average snow and ice cover. The temperature series presented in Fig. 1 suggest that, although there has been a marked drop in temperature in the north-eastern Atlantic, it still has some way to go before attaining values similar to those for the 1920s and there is no evidence from any of the variables that the recent changes are in any way different from those of similar, or even larger, magnitude that have occurred during the past hundred years.

An approximately 10-yr periodicity in sea-surface temperature variations has been reported for several areas of the North Atlantic by Maximov[11] (based on anlyses of data presented by Smed[7] for the period 1900–1960) and also by Southward et al.[12], for a station in the English Channel from the early 1920s. Both reports draw attention to a possible

relationship with the 10–11 yr cycle in the numbers of sunspots.

The sea-surface temperature data for Marsden squares 182 and 145 have been analysed by a filter technique based on the calculation of eigenvectors of the matrix of serial correlations for a specified maximum lag. These analyses showed that the dominant systematic pattern of variation of the residual time series, following the removal of the long term trends contains considerable 10–11-yr periodicity which tends to be in phase with the sunspot cycle, with high temperature anomalies corresponding with sunspot maxima.

The recent work of Parker[13] on pressure distributions in January and July for the period 1750–1958 clearly associates a deepening of the Iceland low in January with the descending phase of the sunspot cycle. This is supported, in general terms, by the data presented by Lawrence[14] of pressure distributions for the month of June for the last eight sunspot cycles (1880–1960). The phase relationships are somewhat confused but the most likely implied association is between a deep Iceland low and positive temperature anomalies. This appears to conflict with the general pattern of intense westerlies and low temperatures. In this case, however, the deepening of the Iceland low is associated with positive pressure anomalies in the area of the Barents Sea and therefore with an increase in southerly anomaly winds over the North Sea and at least some way out over the north-eastern Atlantic. This seems likely to account for the positive relationship between sea-surface temperature and sunspots at the English Channel station and may well apply to areas further out in the Atlantic.

This study is part of the programme of the Institute for Marine Environmental Research, a component of the NERC; it was commissioned in part by the Ministry of Agriculture, Fisheries and Food.

J. M. COLEBROOK

Institute for Marine Environmental Research,
Citadel Road,
Plymouth, UK

1 Brooks, C. E. P., *Geophys. Mem.*, 34 (1926).
2 Bjerknes, J., in *The atmosphere and the Sea in motion* (edit. by Bolin, B.) (New York, 1959)
3 Rodewald, M., *Rapp. cons. Explor. Mer.*, 162, 139–153 (1972).
4 Martin, J H., *Rapp cons. Explor. Mer.*, 162, 159–166 (1972).
5 Namias, J , *Mon. Weath. Rev.*, 97, 346–348 (1969).
6 Cohen, T. J., and Sweetser, E L, *Nature*, 256, 285–286 (1975).
7 Smed, J , *Spec. Publs but Commun. N.W. Atlant. Fish.*, 6, 821–826 (1965).
8 Iselin, C.O'D., *Pap phys. Oceanogr. Mit*, 8, pp. 40 (1940)
9 Schroeder, E. H., *Deep Sea Res*, 12, 323–343 (1965)
10 Wahl, E W., and Bryson, R A., *Nature*, 254, 45–46 (1975)
11 Maximov, I. V , *Rapp cons. Explor. Mer.*, 162, 159–166 (1972).
12 Southward, A J., *et al*, *Nature*, 253, 714–717 (1975).
13 Parker, B. N., *Met. Mag.*, Lond., 105, 33–34 (1976).
14 Lawrence, E. N., *Weather*, 20, 334–343 (1965).

Mantle composition derived from the chemistry of ultramafic lavas

ESTIMATES of the composition of the Earth's mantle have been based on both geochemical and geophysical studies (refs 1–3). As might be expected, there are some discrepancies between the estimates particularly for Na_2O and TiO_2. Such discrepancies in part arise from uncertainties concerning the degrees of melting which many volcanics represent and, possibly, the restricted tectonic settings of the ultramafic nodules studied. High MgO basalts and ultramafic lavas, some of which have been termed komatiites[4] occur among the Archaean volcanic rocks. They represent exceptionally high degrees of melting (up to 50% or more) of peridotite source regions. Potentially such volcanics may be used to make a precise estimate of the composition of their mantle source region. High MgO lavas are well preserved in the Archaean Belingwe greenstone belt of southern Rhodesia, where they form part of a sequence deposited
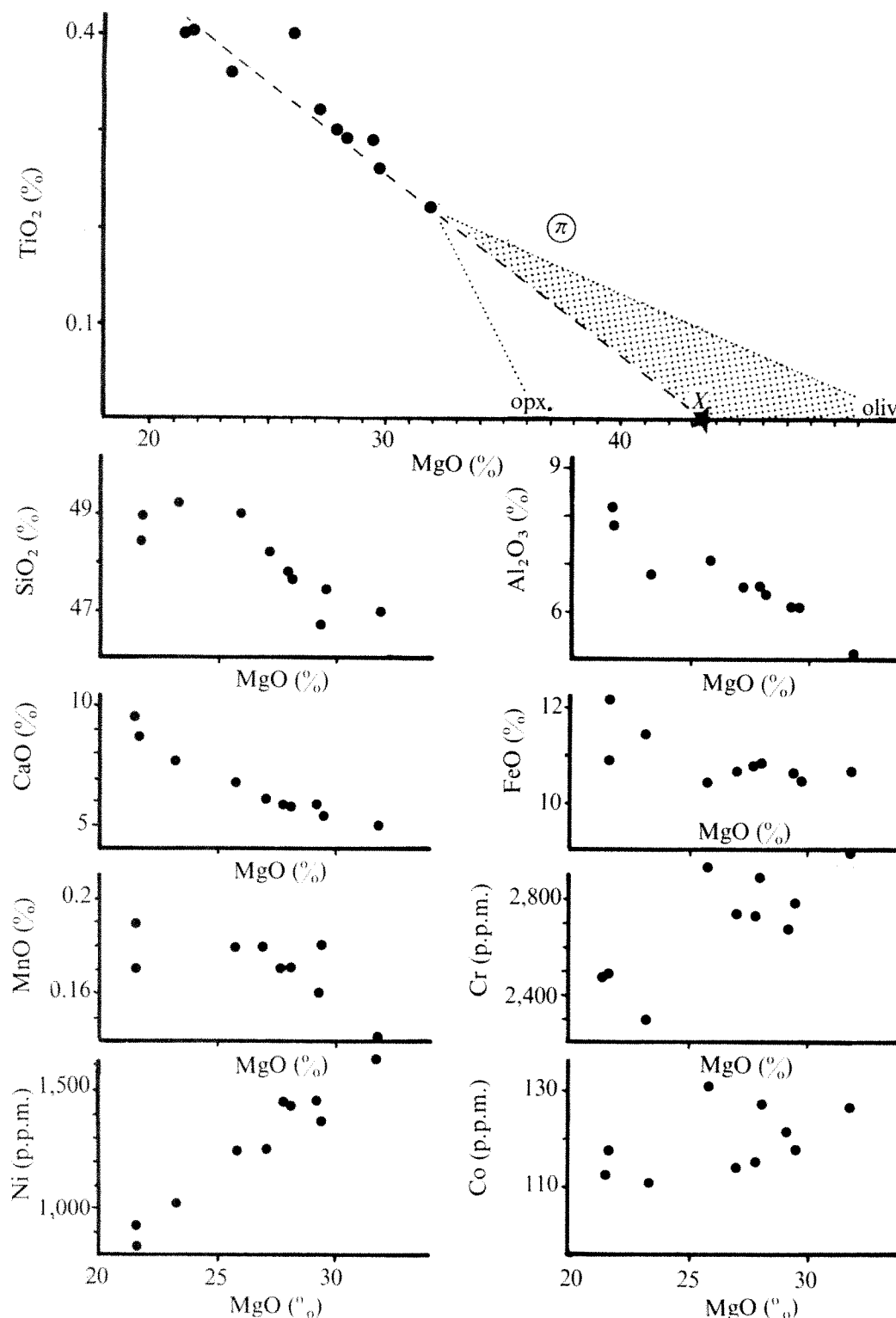
Fig. 1 Major and trace element compositions of ultramafic lavas rescaled to ignore the water content of the rock. The TiO_2 against MgO graph illustrates the method of estimating mantle composition. Regression lines with MgO as the independent variable give liquid compositions at 22% and 32% MgO (Table 1). X is the composition of the solid residue added to the 22% MgO liquid to give the 32% MgO liquid during partial melting

(assuming that TiO_2 is partitioned completely into the liquid). The solid residuum in equilibrium with the 32% MgO liquid is constrained to lie between X and an olivine composition of $Fo_{93.5}$ (see text). π = Archaean pyrolite. The shaded area shows the range of possible source compositions if the residuum evolves towards olivine.

unconformably on granitic crust[5]. We assume that the occurrence of pillow lavas in this sequence containing both skeletal olivine spinifex ($Fo_{92.5}$ in a rock with 28% MgO, E. G. Nisbet, M. J. Bickle and A. Martin, unpublished) and skeletal olivine microphenocrysts indicates that these particular lavas have a composition close to that of the liquid.

Here we present an analysis of these rocks. Only analyses of pillow lavas and thin (2–12 m) flows are considered since rocks with coarse spinifex zones or thicker ultramafic layers are less likely to represent liquid compositions. Because all of these lavas have undergone low greenschist metamorphism (relict clinopyroxene and olivine are preserved)[5] we include only

Table 1 Compositions of various rocks

	Liquid composition at 22% MgO (calculated from Fig. 1)	Liquid composition at 32% MgO	Composition of X	Composition of Fo_{92.8} olivine in a 33% MgO lava^r	Possible source composition if 32% MgO liquid (2) is in equilibrium with a residuum of olivine and orthopyroxene in same ratio as X (olivine Fo_{91}, orthopyroxene En_{91.1})	Possible source composition if 32% MgO liquid is in equilibrium with Fo_{92.8} olivine	Average garnet lherzolite xenoliths in kimberlite[1]
SiO₂ (%)	49.0	46.9	44.5	41.41	45.9	44.8	46.2
Al₂O₃	7.7	5.3	2.7	0.0	2.2	2.7	2.16
Fe₂O₃	1.3	1.1	1.0	0.0	0.4	0.6	
FeO	10.2	9.2	8.1	6.27	8.6	8.1	6.82
MgO	22.0	32.0	43.2	51.6	40.7	40.8	42.0
CaO	8.6	4.5	0.0	0.19	1.7	2.4	1.64
MnO	0.19	0.16	0.11	0.12	0.08	0.11	0.11
Na₂O*				0.0	0.1	0.1	0.17
K₂O*				0.0	0.025	0.025	0.15
P₂O₅*				0.0	0.001	0.001	—
NiO	0.12	0.21	0.32	0.51	-0.16	0.20	0.31
Cr₂O₃	0.4	0.43	0.48	0.18	0.20	0.24	0.35
TiO₂	0.4	0.22	0.0	0.0	0.08	0.11	0.12
Ce (p.p.m.)†	2.4	1.4	0.011	0.014	0.45	0.78	—
Yb†	0.91	0.60	0.017	0.006	0.20	0.33	—

*Values for Na₂O, K₂O and P₂O₅ in columns (5) and (6) are approximated using analyses from thicker but less altered ultramafic units containing euhedral olivines (E. G. Nisbet, M. J. Bickle, and A. Martin, unpublished data).

†Olivine:liquid and orthopyroxene:liquid distribution coefficients of 0.01 and 0.002 for Ce and 0.01 and 0.1 for Yb (ref. 10), are used to calculate their concentrations in columns (3) and (4).

those elements which are relatively unaffected by low grade metamorphism. REE patterns are considered not to have been significantly altered during metamorphism if they are smooth (without any marked kinks or unevenness) and parallel in different samples of similar rock types.

The major and trace element composition of ten ultramafic lavas are plotted against their MgO contents in Fig. 1. Like some other Archaean ultramafic lavas, these have low incompatible element contents and CaO–Al₂O₃ ratios around unity. With the exception of Na₂O and K₂O all the variation diagrams show reasonable linear correlations (Fig. 1). The correlations permit an estimate to be made of the bulk composition of the material added to the melt (in a partial melting model) or removed from the melt (crystallisation model) as it progresses from 32 to 22% MgO. It is important to note that this material (henceforth termed X) need never have existed as a real solid. Its MgO content is taken to be that at TiO₂ = 0.0, since only olivine and orthopyroxene are likely to be in equilibrium with such magnesian liquids. The normative composition of X does in fact contain olivine and a significant amount of orthopyroxene. A similar MgO content is obtained for CaO = 0.0. Titanium was chosen as an incompatible element likely to be almost exclusively partitioned into basic melts[4] (minerals from harzburgite nodules have a very low TiO₂ content).

The chemical variation in Fig. 1 suggests that the melts must have evolved leaving a residuum containing both olivine and orthopyroxene. Moreover, the graphs of Cr₂O₃ and Al₂O₃ against MgO indicate that reasonable initial mantle compositions of these phases cannot control these trends. We consider, therefore, that the melts evolved in equilibrium with a solid residuum containing high NiO olivine and high Cr₂O₃, Al₂O₃ orthopyroxene.

Some constraints may be placed on the likely composition of the source by attempting to sum the compositions of the melt and residual phases in their correct proportions.

Two estimates must be made: first, the MgO content of the source, from which the degree of partial melt needed to produce the 32% MgO liquid can be calculated. We arbitrarily chose MgO 41% (based on published analyses of ultramafic nodules), Second, we must estimate the composition of the residuum in equilibrium with the most magnesian liquid observed at the

surface (32% MgO). The Mg–Fe ratio of the residuum is likely to be greater than that of X.

Estimates of the source composition have been made for two cases; an olivine : orthopyroxene residuum in the same ratio as the normative content of X; and a pure olivine residuum (similar to the Fo_{92.8} olivine observed in an ultramafic lava^r). Demonstrably neither of these situations can be correct but the actual source composition probably lay within these limits.

The computed major element contents are remarkably similar to that of garnet lherzolite nodules in kimberlites (Table 1). The total Fe and Al₂O₃ are slightly higher (Fe is dependent on the Mg:Fe ratio assumed for the residuum) and SiO₂ possibly lower than the average of lherzolite xenoliths (Table 1). If a mantle MgO content of 37% were assumed, the SiO₂ content would be higher and the TiO₂ and CaO content lower than pyrolite[3].

The concentration of the incompatible elements (TiO₂, CaO, K₂O, NaO, REE and to a lesser extent Al₂O₃ and MnO) in the source depend only slightly on the concentration of these elements assumed for the residuum (see Table 1). The calculated concentrations of compatible trace elements (NiO, Cr₂O₃) is, in contrast, strongly dependent on the composition assumed for the residuum.

The REE patterns of the komatiites exhibit a depletion in light (relative to heavy) rare earths (Fig. 2). The Ce and Yb contents of the source are reasonably well constrained between 0.51–0.78 and 0.23–0.33 p.p.m. respectively. These estimates are less than the average chondrite value for Ce and between 1 and 1.5 times chondrite for Yb and are significantly lower than previous estimates of REE abundances in the mantle (3–5 times chondrite, see ref. 8). This particular mantle segment was probably less depleted in light REE than the source of some modern mid-ocean ridge tholeiites[9].

Our work shows that a source mantle similar in composition to modern garnet lherzolite nodules in kimberlites could have produced komatiitic liquids. A pyrolitic source, even when extrapolated to similar MgO contents, is enriched in TiO₂ and CaO and depleted in SiO₂ relative to the source of these komatiites. Finally, if the approximately chondritic value of REE abundance in the komatiite source is typical of large portions of the upper mantle, then the degree of total REE enrichment
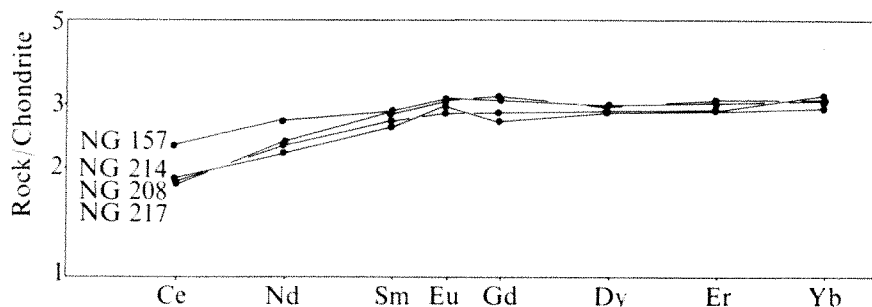
Fig. 2 Chondrite normalised rare earth distribution patterns from ultramafic lavas (23–30% MgO) Belingwe, Rhodesia. Analysis carried out by the mass spectrometric isotope dilution technique described by Hooker *et al.*[11].

(relative to source) observed in both Archaean tholeiites and mid ocean ridge basalts may be much greater than had previously been imagined[8,12].

We thank Pam Shreeve and Mike Hepher for their help. This work was supported by the NERC and the Archaean Crustal Studies Project of the University of Rhodesia.

M. J. BICKLE
C. J. HAWKESWORTH

Department of Earth Sciences,
The University,
Leeds LS2 9JT, UK

A. MARTIN

Geological Survey of Rhodesia,
Salisbury, Rhodesia

E. G. NISBET

Department of Geology,
Parks Road, Oxford, UK

R. K. O'NIONS

Lamont Geological Observatory,
Palisades,
New York 10964

1 Harris, P. G., Hutchinson, R., and Paul, D. K., *Phil. Trans. R. Soc.*, A271, 313 (1966).
2 Ringwood, A. E., *Advances in Earth Sciences*, 287 (Cambridge, Massachussetts, 1966).
3 Green, D. H., *Geology*, 3, 15 (1975).
4 Viljoen, M. J., and Viljoen, R. P., *Spec. Publ. geol. Soc. S. Africa*, 2, 55 (1969).
5 Bickle, M. J., Martin, A., and Nisbet, E. G., *Earth planet. Sci. Lett.*, 27, 155 (1975).
6 Watson, E. B., *Contrib. Miner. Petrol.*, 56, 119 (1976).
7 Green, D. H., Nicholls, I. A., Viljoen, M., and Viljoen, R. P., *Geology*, 3, 11 (1975).
8 Gast, P. W., *Geochim. cosmochim. Acta*, 32, 1057 (1968).
9 Hawkesworth, C. J., and O'Nions, R. K., *J. Petrol.* (in the press).
10 Schnetzler, C. C., and Philpotts, J. A., *Geochim. cosmochim. Acta*, 34, 331 (1970).
11 Hooker, P. J., O'Nions, R. K., and Pankhurst, R. J., *Chem. Geol.*, 16, 181 (1975).
12 Arth, J. G., and Hanson, G. N., *Geochim. cosmochim. Acta*, 39, 325 (1975).

Photochemical ozone in the atmosphere of Greater London

ATMOSPHERIC pollution in Greater London is usually associated with high concentrations of smoke and sulphur dioxide, but continual changes in human activity can lead to other types of pollution. One example is the reported occurrence of elevated ozone concentrations indicative of photochemical activity[1]. I present details of the spatial and temporal variation of ozone concentrations experienced in Greater London during the summer of 1975. The data are presented, in part, with reference to the Greater London Council (GLC) guideline concentration for ozone of 0.08 p.p.m. (1-h average)[2,3]. Observations indicate that photochemical ozone in and around London is likely to have a dual origin, namely, the European continent (see ref. 4), and photochemical reactions involving locally emitted precursor pollutants.

During the period May–September 1975 the Scientific Branch of the GLC monitored ozone at three sites in Greater London, two of which were located in suburban areas at Teddington and Hainault, with the third at County Hall in the centre. Teddington lies 18 km south-west and Hainault 19 km north-east of the centre. The suburban sites were equipped with ozone detectors operating on the principle of ultraviolet absorption. At County Hall a chemiluminescent ozone analyser was used. All three sites were as far removed as possible from local sources of pollution, especially heavily-trafficked roads. The ozone instruments were cross checked and calibrated at the beginning and end of the season against a constant ozone source, itself standardised against a neutral-buffered KI solution[5]. Further cross checks were made between the two types of instrument in polluted atmospheres. Ozone concentrations were recorded at 30-s intervals on magnetic tape suitable for computer processing.

Figure 1 shows maximum hourly mean concentrations of ozone at each site during the monitoring period. Occurrences of elevated ozone concentrations are superimposed on a background concentration of 0.02–0.04 p.p.m. The background seems to represent the natural ozone content of the troposphere in this region, and is consistent with measurements reported for other parts of Britain[4]. With reference to the GLC guideline concentration the first conclusions that can be drawn are:

Fig. 1 Daily maximum hourly means of ozone: *a*, Hainault; *b*, County Hall; *c*, Teddington.

April 30   May 30   June 29   July 29   August 28   September 27
1975
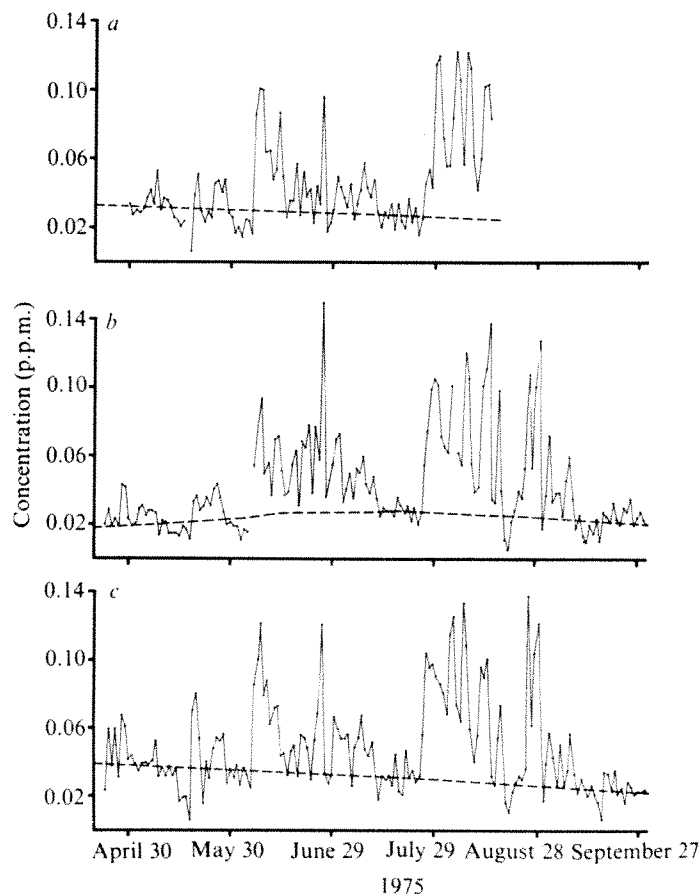
● During May to September 1975 ozone concentrations exceeded the GLC guideline on approximately 15% of days.
● The guideline was exceeded for approximately 100 h at Hainault, 79 h at County Hall and 129 h at Teddington during this period.
● The highest hourly mean concentration recorded at these sites was 0.15 p.p.m. at County Hall, on June 26, 1975.

A study of the meteorological records of the London Weather Centre shows that days with high afternoon ozone concentrations generally coincide with both low early morning wind speeds, usually of less than 3 m s⁻¹, and high solar radiation[*]. An association of oxidant incidents with inversion conditions as measured at Cardington, Bedford-shire, has also been found; it is statistically significant at the 0.1% level. During the summer of 1975 the GLC guide-line was exceeded at one or more of the Greater London sites on 56% of occasions on which the 0500 Balthum ascent from Cardington registered an inversion strength of 4 °C or more. Conversely, for weaker or nonexistent inversions, the guideline was exceeded on only 4% of occa-sions. Figure 2 shows such information for the period June 25–27, 1975. The key factor in the ozone incident of June 26 is thought to be the development of a strong inversion during the previous night, which encouraged the buildup of locally generated oxidant precursor pollutants, oxides of nitrogen and hydrocarbons. These subsequently became in-volved in photochemical reactions during the sunlight hours of June 26. The sequence was not repeated the following day, when overnight cloud cover inhibited inversion forma-tion.

Isolated occurrences of high ozone concentrations, as on June 26 were, however, uncommon. Generally, days with high ozone concentrations tend not to occur singly, but over periods of several days or more[*], and Greater London was no exception during 1975, with the main oxidant incidents

Table 1 Comparison of wind direction (1100–1500) and ozone concentrations at Teddington and Hainault

Date (1976)	Maximum hourly mean	Ozone concentration (p.p.m)	Wind direction (degrees)
	Teddington	Hainault	
June 8	0.122	0.100	060
June 10	0.088	0.065	020
June 13	0.073	0.087	290
July 27	0.104	0.054	120
July 31	0.087	0.072	080
August 1	0.081	0.056	060
August 3	0.116	0.084	120
August 13	0.090	0.103	190

occurring during June 7–12 and July 27–August 14. During the latter period the guideline was exceeded on 14 of the 19 days.

A study of the appropriate Daily Weather Reports shows that on all days with high oxidant levels the synoptic con-ditions were dominated by slow moving anticyclonic cells

Fig. 2 Ozone and meteorological data for the period June 25–27, 1975: a, Hainault; b, County Hall; c, Teddington.
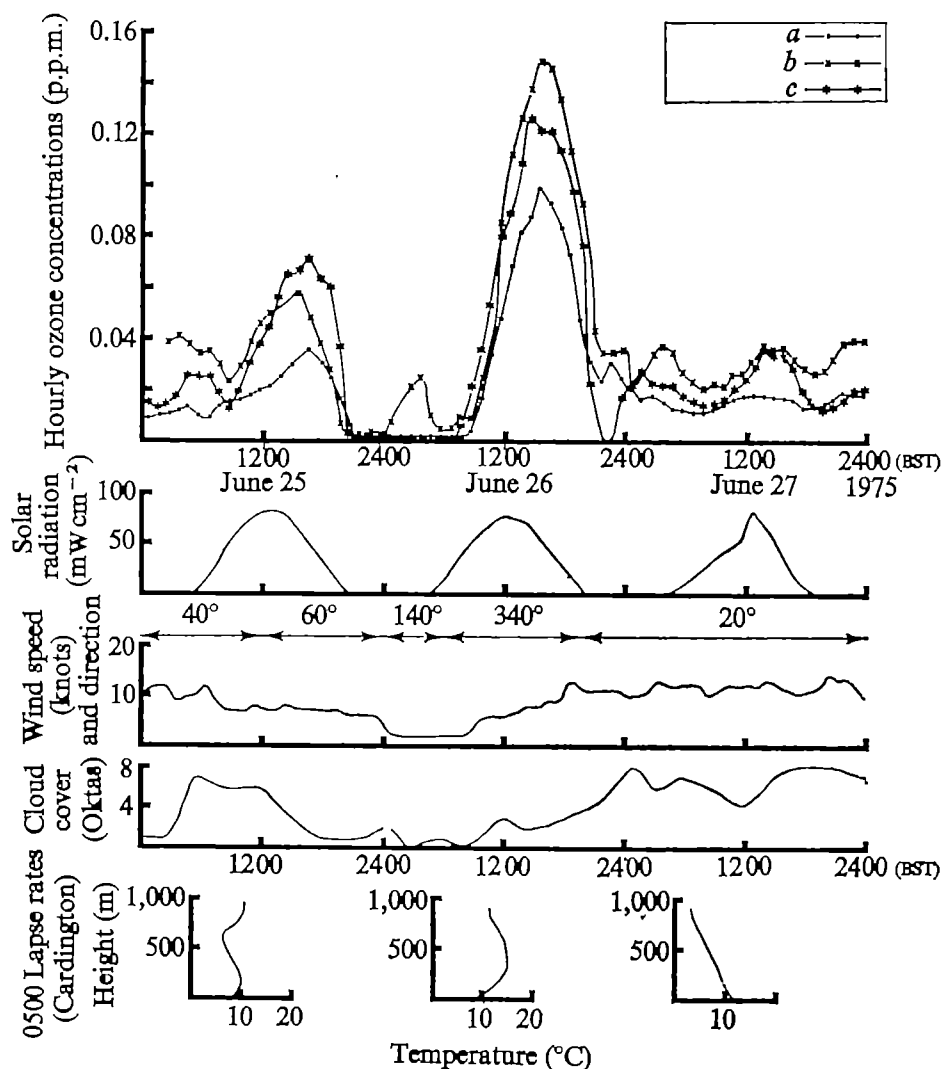
and/or weak synoptic pressure gradients. Anticyclonic conditions are generally characterised by light winds or calm conditions with clear skies. The latter lead to nocturnal inversion, which in turn encourages the buildup of precursor pollutants throughout the whole anticyclonic cell, and the consequent production of oxidants over large areas at any one time. Because the anticyclones are slow moving the conditions tend to persist for days at a time. It is also apparent that on most occasions these cells were centred on the North Sea, North Germany, or Scandinavia. Thus, southern Britain, including the Greater London area, would have been experiencing a light and predominantly easterly air stream of continental origin. This accords with the observations of Cox *et al.*[4], which indicate that photochemical ozone and its precursors may be transported into Britain from continental Europe. In fact, examination of air mass trajectories for 1975 by using the geostrophic wind, as calculated from the Daily Weather Report, shows that most air masses over Greater London on high oxidant days were of continental origin[6].

But a close inspection of the relative values of the maximum hourly means at the Teddington and Hainault sites on high oxidant days on which there was a significant difference in concentration between the two locations shows that the higher concentration is manifest at the site for which the incoming air mass had had the longest trajectory over Greater London (Table 1). In addition, on days characterised by exceptionally long periods with low wind speeds, such as June 26, the highest measured concentration in London, and indeed in the whole of Britain (H.N.M. Stewart, personal communication), occurred in the city centre.

The most likely explanation for these phenomena is that emissions of precursor pollutants from Greater London are reacting photochemically to produce oxidants. The locally generated precursors drift downwind during the late morning and early afternoon and produce higher ozone concentrations on the downwind side, and possibly higher concentrations still beyond the GLC boundary. On particularly calm days the precursors do not drift as far and maximum ozone concentrations tend to arise close to the maximum density of emissions, which is probably in central London. This ties in with the earlier mentioned observation that the GLC guideline was exceeded more in west than in east London, because local winds have tended to be from the eastern sector on high oxidant days.

The picture which emerges is of two sources of oxidants contributing significantly to high ozone concentrations in the atmosphere of Greater London. First, during anticyclonic weather in northern Europe precursor pollutants collect over large areas in what virtually amounts to an outdoor smog chamber. Widespread elevated ozone concentrations are thus formed more or less simultaneously. Ozone, which is rapidly destroyed at the Earth's surface, is then able to persist for several days because low level nocturnal inversions, which form frequently under anticyclonic conditions, inhibit downwind mixing. This allows its transportation over distances of perhaps 1,000 km or more[4], with a possible influx to Greater London. Second, the same synoptic weather conditions favour the simultaneous buildup of locally generated precursor pollutants, and the subsequent formation of oxidants in London's atmosphere.

With forecasts that road traffic will increase significantly in the next decade, continued vigilance is necessary so as to detect, at an early stage, any deterioration in the situation. Although it is apparent that corrective action, should it be required, must be taken in the context of Europe as a whole, the problem of local generation of oxidants warrants individual study as the optimum abatement strategy may vary from one region to another.

I thank R. Hume for assistance with monitoring and Drs

Environmental Sciences Group,
Scientific Branch,
Greater London Council,
County Hall,
London SE1 7PB, UK

1 Derwent, R. G., and Stewart, H. N. M., *Nature*, **241**, 342–343 (1973).
2 Greater London Council Minutes, *London's Air-Guideline Concentrations*, 510–511 (July 22, 1975).
3 Reed, S. B., *J. R. Soc. Hlth* (in the press).
4 Cox, R. A., Eggleton, A. E. J., Derwent, R. G., Lovelock, J. E., and Pack, D. H., *Nature*, **255**, 118–121 (1975).
5 *Selected Methods for the Measurement of Air Pollutants*, United States Public Health Service Document No. 999-AP-11 (US Government Printing Office, Washington, DC, 1965).
6 *Greater London Council Research Memorandum No. 485* (Greater London Council Bookshop, The County Hall, London, 1976).
7 Bilger, R. W., Australian Report Part I, *Oxidants and their Precursors in the Atmosphere* (Department of the Environment and Conservation, Canberra 1974).

Ozone levels in central London

MEASUREMENTS of ozone and some primary pollutants, including oxides of nitrogen and total hydrocarbons, have been in progress at the Central London background site of this laboratory since July 1972. A preliminary note[1] discussed the early results from this site; here we present a statistical analysis of the relationship between ozone levels in Central London and some meteorological parameters measured at the London Weather Centre, ~0.4 km from the sampling site. The data were collected during the periods July–September 1972 and April–September in both 1973 and 1974.

The sampling point at the site, which is located at Endell Street, Holborn, London, is ~10 m above ground, well removed from traffic or any particular pollutant source and is not unrepresentative of the background air in Central London.

An analysis of the time of occurrence of the ozone maximum hourly mean concentration for all the days considered is summarised in Fig. 1. This conforms to the general pattern found in other studies[1,2] with a small percentage of maxima occurring in the earlier hours of the day, arising from the diffusion of stratospheric ozone and the majority of maxima, from photochemical activity, occurring between 1500 and 1600 LT. It is interesting to note that relatively high (>10.9 parts per 10[8]) levels were observed late in the day (1900–2100 LT) on some occasions, a fact which has been taken by Steinberger and Balmor[3] to be indicative of non-local production of ozone.

Fig. 1 Time of occurrence of the maximum hourly mean ozone concentration at Endell St, London for the summer periods 1972–1974.
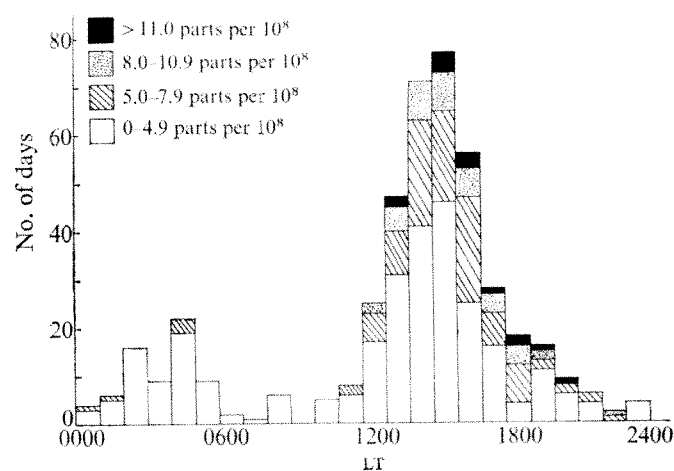
Table 1 Analysis of the dependence of elevated ozone levels on wind direction

	Wind direction							
	N*	NE	E	SE	S	SW	W	NW
No. of days in given direction	51	67	51	21	40	80	72	38
% of total	12	16	12	5	10	19	17	9
No. of days with $[O_3] \geqslant 8.0$ p.p.10^8 in given direction	4	13	11	5	9	6	6	3
No. of days with $[O_3] \geqslant 8.0$ p.p.10^8 as % of total in given direction (f)	8	19	22	24	23	8	8	8

*Any direction from 337.5°–22.5° is taken as N, 22.5°–67.5° as NE and so on.

An attempt was made to correlate the daily ozone maximum hourly mean concentration with average values for the whole day and the period 0600–0900 h for the primary pollutants NO, NO$_x$ and total hydrocarbons. No significant relationship was found.

A preliminary bivariate analysis of the ozone levels and selected meteorological parameters indicated that the most important factors were maximum daily temperature, relative humidity and daily insolation. A stepwise multiple regression analysis was then performed using the logarithms of these

Fig. 2 Plot of values of ozone concentrations calculated from equation (2) against observed values.

parameters as independent variables with the logarithm of the ozone concentration as dependent variable. The number of days for which valid data were available was 453. The regression equation obtained was

$$[O_3] = 0.60 \, T^{0.64} \, I^{0.31} \, H^{-0.43} \tag{1}$$

where T is the maximum daily temperature (°C), I is the total daily insolation (mWh cm^{-2}) and H is the relative humidity (%). The multiple correlation coefficient obtained was 0.623. Adding wind speed to the regression gave no significant improvement. An analysis of the residuals yielded a Durbin–Watson statistic of 1.145 indicating autocorrelation in the ozone levels. On including the logarithm of the previous day's ozone concentration in the set of independent variables the multiple

correlation coefficient improved to 0.712. The regression equation obtained was, in an obvious notation

$$[O_3]_t = 0.68 \, [O_3]_{t-1}^{0.38} \, T_t^{0.37} \, I_t^{0.28} \, H_t^{-0.36} \tag{2}$$

The partial correlation coefficients of the independent variables are, in the order in which they appear in equation (2), 0.44, 0.19, 0.24 and −0.17, (all significant at the 1% level).

The standard error in reproducing observed ozone levels using this model is ±2.1 parts per 10^8 (see Fig. 2) so that in predicting elevated levels, that is, those $\geqslant 8.0$ parts per 10^8 this implies a confidence limit of ±~26%. We therefore present equation (2) as a model for ozone concentrations at the Central London background site.

We have attempted to analyse the effect of wind direction as measured at the London Weather Centre on the ozone levels. The results of this analysis are shown in Table 1 and a plot of the percentage of days with wind from a given direction when levels $\geqslant 8.0$ parts per 10^8 occurred is shown in Fig. 3. It is clear from this histogram that the majority of elevated zone levels occur on days when the wind is from a direction ranging from north-east through to south. These conditions prevail, in general, when an anticyclone is present over the North Sea or north-western Europe when the air flow is from the Continent.

It should be borne in mind that the wind direction as measured at a particular location, in our case the London Weather Centre,

Fig. 3 Plot of number of days with ozone concentration $\geqslant 8.0$ parts per 10^8 expressed as a percentage (f) of the total number of days with wind in a given direction.
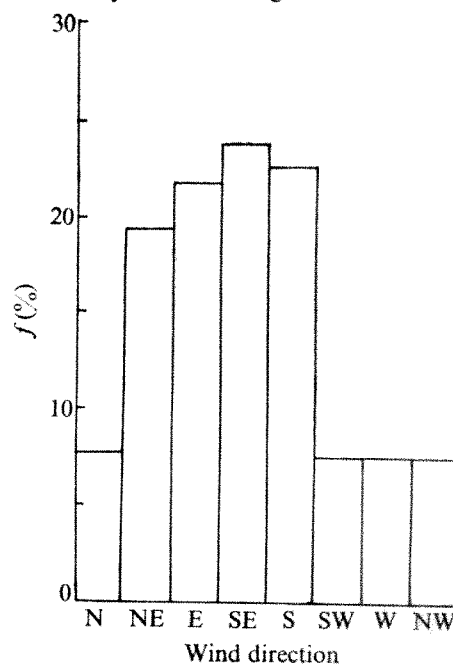
is not necessarily a precise indicator of the actual trajectory of the air mass which might be responsible for the transport of pollutants. Consequently further refinement of the model to include wind direction in a more quantitative manner, by stratifying the data in wind sectors for example, was not felt to be justified.

From a consideration of the wind direction analysis and the times of occurrence of the elevated ozone levels it is likely that a significant proportion of the ozone measured at Endell Street is transported there, but further investigation would be necessary to determine the source areas. One possibility is that the continental air mass contains precursor pollutants and produces ozone as it flows over north-western Europe giving rise to elevated levels even at rural sites well away from sources of pollution. Elevated ozone levels measured at Adrigole on the West Coast of Ireland lend support to this idea and Cox *et al.*[2] have discussed this point at length, concluding that ozone and precursors could be transported over distances up to 100–1,000 km.

H. N. M. STEWART
E. J. SULLIVAN
M. L. WILLIAMS

Department of Industry,
Warren Spring Laboratory,
PO Box 20, Gunnels Wood Rd,
Stevenage, Herts SU1 2BX, UK

[1] Derwent, R. G., and Stewart, H. N. M., *Nature*, **241**, 342–343 (1973).
[2] Cox, R. A., Eggleton, A. E. J., Derwent, R. G., Lovelock, J. E., and Pack, D. H., *Nature*, **255**, 118–121 (1975).
[3] Steinberger, E. H., and Balmor, Y., *Nature*, **241**, 341–342, (1973).

Gravity counteracts light-induced inhibition of root growth

THE view that roots are generally indifferent to light has recently been revised in the face of evidence that light can be inhibitory to the growth of some roots[1-10]. The discovery[6,11] that light affects the production of an inhibitor prompted the suggestion[8] that sensitivity to light in roots is related to the system controlling geotropic response, which is also dependent on the presence of an inhibitor[12,13]. Since the onset of geotropic responsiveness in some roots is dependent on exposure to light[14-16], the suggestion seems well founded and attempts are being made to clarify the extent to which the mechanisms, triggered when light or gravity affects the root, are interrelated. The investigation described here was prompted by the report[9] that roots of cress seedlings grown in Petri dishes in the light were stunted in comparison with similar seedlings in the dark. If light promotes geosensitivity, distorted growth could develop in a situation where the root is unable to realign itself, as for example when it is restricted to the horizontal plane. A study was therefore made of the influence of the plane of growth on the light-induced inhibition of root

elongation in cress seedlings, from which it seems that although light does not enhance geosensitivity, light-induced inhibition is manifest only in horizontally grown roots and can be prevented by subjecting the roots to a force field.

Seeds of *Lepidium sativum* L. cv. Curled were moistened with 0.1 mM $CaCl_2$, pH 6.5, and left for 20 h at room temperature in normal lighting conditions. Seeds which were at a uniform stage of germination—the testa having split open but before the emergence of the radicle—were then placed with the radicles pointing downwards on a 11.5-cm circle of Whatman No. 1 filter paper in a glass Petri dish (115×18 mm) containing 6 ml of 0.1 mM $CaCl_2$. For dark treatments Petri dishes were either sprayed externally with black paint or wrapped in aluminium foil. The dishes were placed on a bench continuously illuminated by a bank of fluorescent tubes giving an incident light intensity of about 20 W m^{-2} at the level of the Petri dishes. The dishes were placed horizontally, parallel to the light source, or at angles of 10°, 45° and near vertical (85° to 90°). For the angled dishes side illumination was also provided but growth showed no dependence on the direction of the light source. Cress roots are not negatively phototropic. Root lengths were measured at 24-h intervals after treatment began, that is approximately 44, 68 and 92 h after the seed was wetted.

Table 1 shows a comparison of root growth of seedlings in the light and in the dark as affected by the plane of growth. The reported inhibition[5] of root growth in the light was evident in horizontal roots at 44 h, and this inhibition continued into the following 2 d, although with time the growth rate of horizontal roots in the light increased. With vertically orientated seedlings in the light there was no inhibition of root growth relative to horizontal roots in the dark. The ability of light-exposed roots to overcome light inhibition is clearly a function of the plane of growth, an angle of 10° having a discernible effect. At 45° the growth rate was not far short of that obtained on the vertical. Since the component of the Earth's gravitational field in the plane of growth is determined by the sine of the angle subtended with the horizontal, an angle of 45° would produce a force of approximately 0.7*g* and therefore it would be expected that this slope would give a growth rate approaching that of the vertical. Roots growing in the dark also showed a positive response to being aligned with the Earth's gravitational field but the percentage increase was much less than in the light.

These results indicate that when the force of gravity acts on a seedling growing in a vertical plane the light-induced retardation of growth observed in the horizontal plane is largely eliminated. Two further questions are thus raised. First, is horizontal root growth inhibited in the light in the absence of gravity and second, is horizontal root growth inhibited in the light when the root is subjected to a force field applied on the horizontal plane? An answer to the first question was sought by placing imbibed seed on moist filter paper on the base of a crystallising dish which was illuminated and, with the base vertical, was then rotated

Table 1 Effect of the plane of growth on the length (mm) of cress seedling roots in the light and dark

Treatment	Time of measurement (h)	Angle subtended with the horizontal			
		0°	10°	45°	85°
Light	44	6.1 ± 0.28	10.5 ± 0.39	12.8 ± 0.40	16.1 ± 0.49
	68	15.7 ± 0.66	35.3 ± 1.73	43.2 ± 0.90	50.5 ± 1.06
	92	30.5 ± 0.91	69.4 ± 2.82	76.8 ± 1.62	90.7 ± 1.86
Dark	44	14.7 ± 0.55	17.1 ± 0.63	20.3 ± 0.51	21.3 ± 0.46
	68	39.5 ± 1.42	51.5 ± 1.31	56.0 ± 1.35	60.2 ± 0.71
	92	65.6 ± 2.44	79.7 ± 2.24	80.3 ± 1.64	88.6 ± 1.99

Twenty seeds were placed in Petri dishes as described in the text and the growth of the roots was measured over 4 d at 25 °C. The Petri dishes were clamped at the angles shown and care taken to ensure uniform illumination at an intensity of approximately 20 W m^{-2}. Each treatment was in duplicate and the values shown are the mean ± s.e. of 40 seedlings.

round a horizontal axis in the manner of a clinostat. The second question was answered by placing a circle of imbibed seed so as to form a hub on filter paper in a Petri dish which was then illuminated and spun at 150–200 r.p.m. (1–2*g*) in a horizontal plane. Table 2 shows that where the effect of gravity was neutralised, and with the root free from the restraints imposed by a horizontal barrier, a light-induced inhibition of root growth was still evident. But this inhibition could be overcome on the horizontal plane by applying a centrifugal force comparable with that of gravity.

The question posed above concerning the geosensitivity of cress roots in light and dark was answered by making a 90° rotation of Petri dishes containing seedlings growing vertically in the light and in the dark. In both cases the roots reorientated themselves with the field of gravity irrespective of whether they were, or had been, in the light or in the dark. The light-induced root inhibition exhibited by horizontally grown seedlings is not therefore attributable to an enhanced geosensitivity not possessed by the seedlings grown in the dark.

Table 2 Effect of force field on the growth of cress seedling roots

Treatment	Radicle length (mm)
Light horizontal	6.7±0.38
As above with centrifugal force	17.8±0.71
Gravity neutralised	
dark seedlings	14.0±0.71
light seedlings	6.2±0.20

Seeds were imbibed for 20 h and then 20 seeds were placed in each of the dishes as described in the text and subjected to treatment for the following 24 h at 25 °C. The centrifugal force applied on the horizontal plane was between 1 and 2*g*. Gravity was neutralised by rotating the vertically-orientated dishes round a horizontal axis at 5 r.p.h. Values shown are mean ± s.e. for each treatment.

It is evident that in cress roots, growth is promoted by the longitudinal force exerted by gravity, one effect of which is to counteract the inhibitory effect of light. Not only root growth is promoted. As the steepness of the plane of growth increases so does the growth of the seedling, the cotyledonous leaves emerging and greening more rapidly than when horizontal. This points to an overall effect of gravity on the rate of hormone production or transport or both.

The cress root is a classic object for the study of amyloplasts, the gravity-sensing organelles present in the root cap[17,18]. Audus[19] has suggested, on the basis of work by Sievers and Volkmann[20], that mechanical pressure exerted by the amyloplasts on the membranes of the endoplasmic reticulum, or some other non-migratory cellular component, could result in the production of a regulatory compound in the root cap. This is envisaged in terms of a differential flow of inhibitor leading to a curvature of the root, but it seems reasonable that a preferred orientation of amyloplasts should also give a minimum flow of inhibitor or an optimum flow of regulators.

It has been suggested that root growth is controlled by two hormonal systems[19,21]. Light is known to promote the acropetal movement of indolyl acetic acid in *Zea* roots[15] and it has been reported that the gravity-promoted basipetal movement of a growth inhibitor in *Zea* is dependent on light[8,22]. The results reported here show that gravity can interact with light antagonistically in the overall control of root growth. Identical results have also been obtained with seedlings of *Lactuca sativa* and *Brassica hirta* in both of which the roots, in contrast to those of cress, are strongly negatively phototropic, and in both of which a pronounced light-inhibition on the horizontal plane is relieved on the vertical plane. This phenomenon might prove useful in the investigation of the changes in hormone production and

transport brought about by amyloplasts sedimenting under the influence of gravity.

I. R. MacDonald

Macaulay Institute for Soil Research,
Craigiebuckler, Aberdeen AB9 2QJ, UK

1 Torrey, J. G., *Pl. Physiol.*, *Lancaster*, 27, 591–602 (1952).
2 Pilet, P.-E., and Went, F. W., *Am. J. Bot.*, 43, 190–198 (1956).
3 Burström, H., *Physiol. Pl.*, 13, 597–615 (1960).
4 Masuda, Y., *Physiol. Pl.*, 15, 780–789 (1962).
5 Wain, R. L., Taylor, H. F., Intarakosit, P., and Shannon, T. G. D., *Nature*, 217, 870–871 (1968).
6 Wilkins, H., Larqué-Saavedra, A., and Wain, R. L., *Nature*, 248, 449–450 (1974).
7 Wilkins, H., Larqué-Saavedra, A., and Wain, R. L., *Ann. appl. Biol.*, 78, 169–177 (1974).
8 Wilkins, H., and Wain, R. L., *Planta*, 121, 1–8 (1974).
9 Larqué-Saavedra, A., Wilkins, H., and Wain, R. L., *Planta*, 126, 269–272 (1975).
10 Tietz, A., *Biochem. Physiol. Pflanzen*, 165, 387–392 (1974).
11 Pilet, P.-E., *C. r. Acad. Sci.*, *Paris*, 276D, 2529–2531 (1973).
12 Gibbons, G. S. B., and Wilkins, M. B., *Nature*, 226, 558–559 (1970).
13 Shaw, S., and Wilkins, M. B., *Planta*, 109, 11–26 (1973).
14 Lake, J. V., and Slack, G., *Nature*, 191, 300–302 (1961).
15 Scott, T. K., and Wilkins, M. B., *Planta*, 87, 249–258 (1969).
16 Tepfer, D. A., and Bonnett, H. T., *Planta*, 106, 311–324 (1972).
17 Zimmermann, W., *Jahrb. wiss Bot.*, 66, 631–677 (1927).
18 Iversen, T.-H., Pedersen, U., and Larsen, P., *Physiol. Pl.*, 21, 811–819 (1968).
19 Audus, L. J., in *The Development and Function of Roots* (edit. by Torrey, J. G., and Clarkson, D. T.), (Academic, New York, 1975).
20 Sievers, A., and Volkmann, D., *Planta*, 102, 160–172 (1972).
21 Wilkins, H., and Wain, R. L., *Planta*, 126, 19–23 (1975).
22 Pilet, P.-E., *Physiol. Pl.*, 33, 94–97 (1975).

Heterochromatin polymorphism and colour pattern in the tiger swallowtail butterfly *Papilio glaucus* L.

THE difference between the black mimetic and yellow non-mimetic female forms of the tiger swallowtail butterfly, *Papilio glaucus* L., is probably controlled by a locus on the Y (W) chromosome (cytoplasmic inheritance is less likely). Since the female is the heterogametic sex in the Lepidoptera black females should produce only black daughters and yellow females only yellow ones, and here we report cytological differences between the two.

It is known that occasional yellow females arise from black ones and vice versa, and in our broods four black females gave rise to a minority of yellow daughters, and two yellow ones to a minority of black[1]. Haldane[2] suggested that these exceptional broods might be explained by occasional crossing over between the X and the Y chromosome, plus a position effect, such that black pigment is only laid down by the appropriate allelomorph when this is on the Y chromosome. Furthermore, he thought a position effect to be particularly likely because of the tendency for inert chromatin to be present on the Y chromosome.

The difficulty with Haldane's explanation of the exceptional *glaucus* broods is the fact that crossing over is believed not to occur in female Lepidoptera[3-5]. Our recent cytological evidence however, lends striking support to Haldane's view that heterochromatin is involved in the expression of the black phenotype.

We have examined some of our *P. glaucus* males and black and yellow females and found that the black females possess and the yellow ones lack the heteropyknotic sex chromatin body (Fig. 1) described by Smith[6] and later shown to be the heterochromatin on the Y (W) chromosome[7]. The yellow butterflies tested were the male and female descendants of two wild yellow females from New Jersey, and the female descendants of a yellow female from West Virginia. All were Smith negative. Among the descendants of three wild black females from West Virginia and one from Chicago, the males were again negative, but in contrast all the females were Smith positive. Thus the black females and the (yellow) males are characteristic of the cytological picture of most Lepidoptera, whereas the yellow females are behaving like those of a minority of species where no heterochromatic body is found in either sex[8]. In the only other minority category known to occur, there is a body in both sexes, and in at least one species (*Papilio machaon* L.) the male is polymorphic for the presence or absence of the body (our work with W. Traut, in preparation).
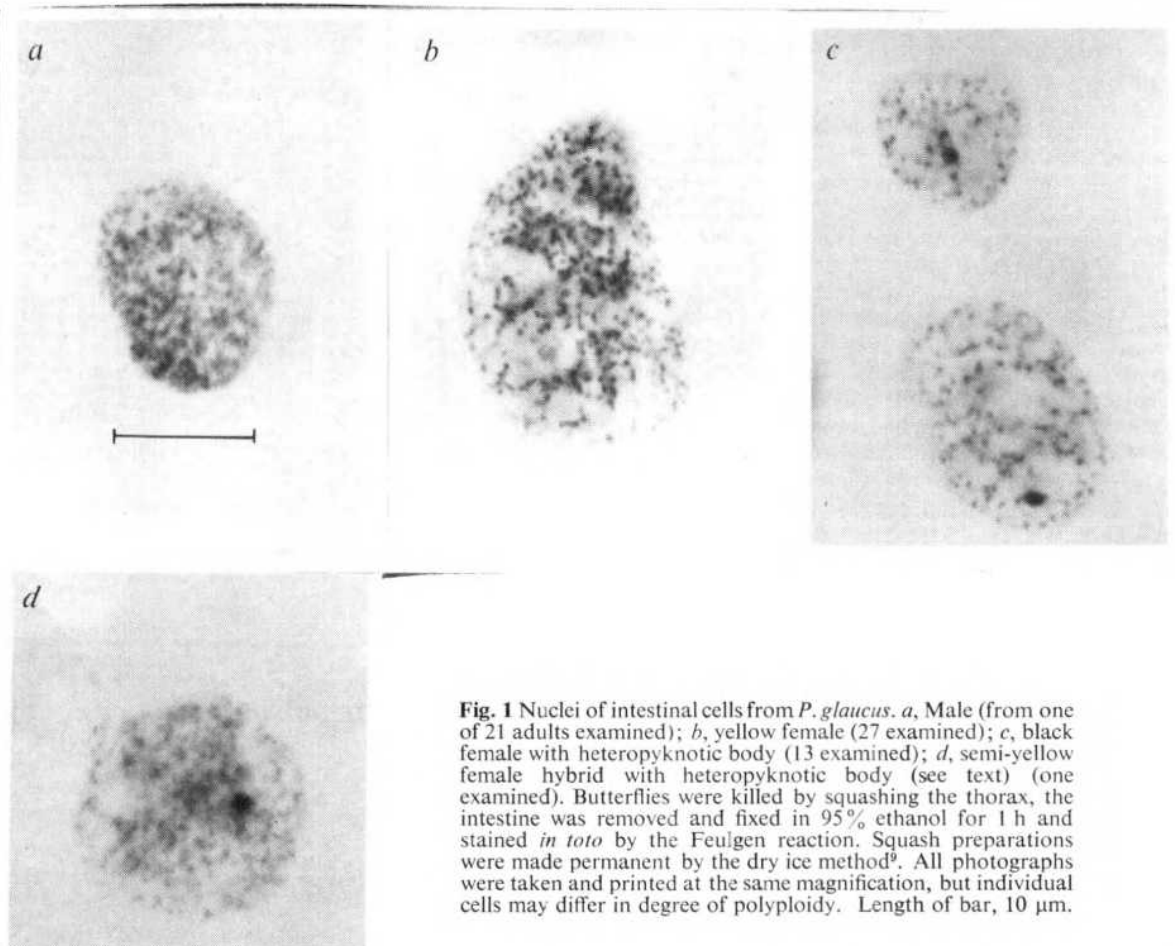
Fig. 1 Nuclei of intestinal cells from *P. glaucus*. *a*, Male (from one of 21 adults examined); *b*, yellow female (27 examined); *c*, black female with heteropyknotic body (13 examined); *d*, semi-yellow female hybrid with heteropyknotic body (see text) (one examined). Butterflies were killed by squashing the thorax, the intestine was removed and fixed in 95% ethanol for 1 h and stained *in toto* by the Feulgen reaction. Squash preparations were made permanent by the dry ice method[9]. All photographs were taken and printed at the same magnification, but individual cells may differ in degree of polyploidy. Length of bar, 10 μm.

P. glaucus is widely distributed in North America, and we have not examined sufficient females from different localities to be certain that the 1:1 correspondence between female colour and heteropyknotic state in our study is not due to geographical heterogeneity in the cytological picture. The correspondence, however, suggests that in the exceptional segregating broods where the mother is black, the Y chromosome bearing the locus for the black gene is occasionally lost from the eggs of black females, who would then produce yellow daughters. It is more difficult to see how the body could be gained as it would have to come from a non-homologous chromosome. Possibly relevant is the fact that, as we noted[1], all our exceptional broods stemmed from females which may well have been related and have a chromosome abnormality resulting in loss or gain of heterochromatin by the Y.

There is another aspect of the Smith body in our work on *P. glaucus*. To be certain that the exceptional females were not mistakes due to the interchange of larvae between broods, we hybridised black *P. glaucus* females with the yellow monomorphic allopatric species *P. rutulus* Lucas because the hybrid larvae are recognisable. This cross produced no adult females, so we backcrossed the male progeny to black *P. glaucus*. We obtained some broods which segregated for black and yellow females and others in which a proportion of the black females had various degrees of yellow scaling, but none was fully yellow. The segregation (25 pure black in 41 females) suggests a 1:1 ratio, a situation quite different from that in pure *P. glaucus*. Since we began to use the Smith technique we have found that the only hybrid semi-yellow female (dissected and found to be a normal female) that has been tested possesses the heterochromatic body characteristic of black *P. glaucus* (Fig. 1*d*): two of her all-black sibs also had the body. In the progeny of the backcross to yellow the body was absent.

The segregation of yellow and black in the *P. rutulus* hybrids seems therefore to be due to an allelomorph suppressing black, introduced from *P. rutulus*. This is supported by the fact that the

segregation in the backcross occurs in the progeny of females that were pure black *P. glaucus*. It seems therefore that the black form can be altered in two ways, either by bringing in a modifier not on the Y chromosome from the allied *P. rutulus*, or by a change on the Y chromosome associated with the absence of the heteropyknotic body within *P. glaucus*.

It is important to know the state of the heteropyknotic body in the three allopatric species *P. rutulus*, *P. multicaudatus* Kirby and *P. eurymedon* Lucas, since the ancestral cytological state is likely to have been that found in black *P. glaucus* today (the majority situation) and it is interesting from the evolutionary point of view to know whether the heteropyknotic body has been lost and regained, or merely lost in yellow *P. glaucus*.

Since our manuscript was submitted adult females have been obtained from black *P. glaucus* females and *P. rutulus* Lucas. On July 8, 1976 two perfect females were made to eclose by Dr Axel Willig of the University of Ulm, using α-ecdysterone. The insects resemble the yellower intermediate backcross females and have a Smith body.

C. A. CLARKE
P. M. SHEPPARD

Department of Genetics,
University of Liverpool,
PO Box 147, Liverpool L69 3BX, UK

URSULA MITTWOCH

Galton Laboratory,
University College,
London WC1 6BT, UK

1 Clarke, C. A., and Sheppard, P. M., *J. Genet*, **56**, 236-360 (1959).
2 in Clarke, C. A., and Sheppard, P. M., *J. Genet.*, **56**, 236-360 (1959).
3 Tazima, Y., *The Genetics of the Silkworm*, (Logos, London, 1964).
4 Suomalainen, E., Cook, L. M., and Turner, J. R. G., *Hereditas*, **74**, 302-304 (1973).

5 Turner, J. R. G., and Sheppard, P. M., *Heredity*, **34**, 265-269 (1975).
6 Smith, S. G., *Sci. Agric.*, **25**, 566-571 (1945).
7 Traut, W., and Rathjens. B., *Chromosoma*, **41**, 437-446 (1973).
8 Traut, W., and Mosbacher, G. C., *Chromosoma*, **25**, 343-356 (1968).
9 Conger, A. D., and Fairchild, L. M., *Stain Technol.*, **28**, 281-283 (1953).

Diet influences attractiveness of urine in guinea pigs

INTRASPECIES chemical communication with urine-borne substances is widespread in mammals. Numerous reports have indicated that the nature of the messages transmitted in urine depends on such factors as the hormonal state, the age, the individual identity and the level of stress of the sender[1,2]. I report here that information transmitted in urinary signals also depends on the diet of the sender. These findings have important implications for the study of animal chemical communication systems[3].

Chemical signals have an important role in social interactions of the guinea pig. In seminatural settings, subordinate females spray males and other females with urine whereas males often spray females during courtship. Scent-marking behaviour in this species results in placing urine as well as glandular secretions throughout the environment[4,5]. In laboratory studies, males spend more time investigating female conspecific urine than male conspecific urine, urine from other species, or water[6,7].

To examine whether the diet of an animal influences the information contained in the urine, 12 individually housed, sexually mature, socially experienced male guinea pigs were given a series of preferences tests (single-sample tests and two-choice tests) between urine collected from donors on different diets. The tester was always unaware of the source of the urine being tested. A finding that urine from one source is preferred to urine from a second source is evidence that the two samples possess different information. The urine to be tested was collected from adult, intact sexually non-receptive female and adult male guinea pig donors using metabolism funnels. Different donors were used in each experiment. All urine was frozen until shortly before use. In the first experiment, urine from four male and four female donors was used. Twenty-two hours before the first urine collection two female donors and two male donors, randomly selected, were provided with crushed (Purina) rat food (RF) and water. The other donors remained on (Wayne) guinea pig food (GPF). All donors' food was removed and urine was collected during the next 2 h. Urine from donors eating the same food (GPF or RF) and from donors of the same sex (♀♀ or ♂♂) was pooled and two series of single sample tests[6,7], separated by 2-d rest, were conducted. Each male was tested with ♀♀ GPF urine, ♂♂ GPF urine, ♀♀ RF urine, ♂♂ RF urine and water. The weights of the donors either increased slightly or remained constant during the 22 h on experimental diets.

The results (Fig. 1) show that females on a RF diet produce urine which is less attractive than that of females eating GPF. This occurred even though the donors were fed the new diet for only 22 h preceding urine collection. Even though the attractiveness of ♀♀

RF urine is, however, depressed, this urine is still investigated more than ♂♂ GPF urine, confirming the previously documented preference for female urine over male urine by male recipients[6,7].

To examine the generality of this finding, urine from female and male guinea pigs fed a semisynthetic diet (Reed-Briggs) was compared with urine from the same animals fed GPF using a simultaneous two-choice method[6,7]. The urine collected when the donors were on a semisynthetic food was less attractive than urine collected when they were on GPF (Table 1, experiments 1 and 2), in agreement with the RF results. When the donors were force fed a standard amount (approximately 16 g over 2 d) of either GPF or semisynthetic diet, the results were similar: GPF urine was
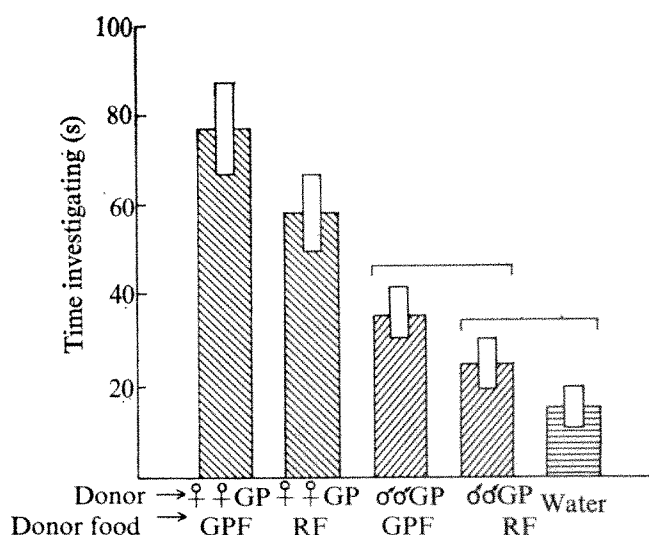
Fig. 1 Results of single-sample preference tests. The heights of the bars represent the mean (± s.e.m.) time (s) that the male guinea pigs have their nose within approximately 1 cm of the urine from four sources and water, each presented singly. The four sources are: ♀♀ GPF, urine that was collected from females eating guinea pig food during the 22 h preceding collection; ♀♀ RF, urine collected from females eating rat food over the same period; ♂♂ GPF, urine collected from males eating guinea pig food, ♂♂ RF, urine collected from males eating rat food; water, charcoal filtered and deionised water. For each substance, each animal had a total of two 2-min periods to investigate a sample (0.2 ml) which was placed in the centre of a clean glass plate (7.5 × 15.0 cm). Each sample was tested individually in the subjects' home cage, one sample each day. The order of presentation was randomised. Overall, there was a significant difference among the samples (represented measures ANOVA, $F[4,44] = 21.98$, $P < 0.001$). Those bars not connected by a horizontal line are significantly different from each other ($P \leq 0.05$, Newman–Keuls test).

Table 1 Mean time (s) investigating each of the two guinea pig urine samples in five preference experiments.

Experiment	Donors' Chow	Time (s)		Choice Donors' Chow	Time (s)	P
1*	Guinea Pig	107.0 ± 18.7	*v*	Semisynthetic	44.1 ± 7.2	<0.01
2*	Guinea Pig	92.2 ± 14.1	*v*	Semisynthetic	36.3 ± 7.1	<0.01
3	Guinea Pig	95.0 ± 23.6	*v*	Semisynthetic	37.4 ± 8.7	<0.01
4	Guinea Pig	108.3 ± 21.2	*v*	Rat	50.8 ± 10.1	<0.01
5a	Guinea Pig	36.9 ± 4.7	*v*	Guinea Pig	30.8 ± 6.9	NS
b	Guinea Pig	57.2 ± 9.5	*v*	Rat	19.2 ± 4.8	<0.01
c	Guinea pig	33.2 ± 8.4	*v*	Guinea pig	33.4 ± 7.6	NS

Each experiment except 5 represents two 4-min choice tests with 12 or 13 subjects where the two samples to be compared were presented simultaneously to a subject; in experiment 5, only one 4-min choice was conducted for each of the three tests. For all experiments the donors were four females except for choice 2, where four males served as donors. Probabilities based on two-tailed *t* tests for related measures. See text for further explanation.
*In choices 1 and 2, urine from donors fed GPF was diluted in the appropriate amount with deionised water before testing, since animals on semisynthetic diet drank more water and produced larger volumes of urine.
NS, not significant.

preferred to urine from donors force fed semisynthetic food (Table 1, experiment 3). One possible explanation for the pattern of results is that abrupt short term alteration of the diet stimulates physiological changes that result in the production and urinary excretion of some substance that the male guinea pig subject finds noxious. To examine this question, two female guinea pigs were raised from birth on crushed RF and water fortified with vitamin C. When the animals were well past sexual maturity, urine was collected and compared with urine from littermates reared on GPF from birth. A different group of 13 males was used as subjects in this test. As before, GPF urine was greatly preferred over RF urine (Table 1, experiment 4).

Different groups of animals served as donors for each experiment described above reducing the possibility that the differences observed could be accounted for by inherent differential attractiveness of urine from different animals. To eliminate completely this possibility, however, a final experiment was conducted. From four previously unused female donors of approximately equal age and weight, two were randomly selected to be on RF and the other two were to remain on GPF. Three pairs of urine samples were collected. Immediately before placing the donor pair on RF, urine was collected for 3 h from both pairs (*a*). Immediately after the one pair had eaten RF for 21 h, urine was similarly collected from both pairs (*b*). Finally, GPF was available to both pairs of donors then for the next 21 h after which urine was collected a third time (*c*). The amount of RF consumed during the 21 h (25 and 28 g for the two donors) was similar to the amount of GPF eaten in the same period by the two control donors (26 and 27 g); similarly, body weight remained constant. All females were sexually unreceptive during the 3-d period. In several previous studies we have failed to find any change in the attractiveness of female guinea pig urine as a function of the stage of oestrus. When the three pairs of urine samples were tested using the simultaneous two-choice procedure, it was found that urine from both pairs of donors which was collected before the diet was changed (experiment 5*a*) and urine which was collected after both pairs of donors had been returned to GPF for 21 h (experiment 5*c*) was equally attractive. Urine collected from donors immediately after one pair had been on RF for 21 h was, however, significantly less attractive than urine from donors on GPF (experiment 5*b*).

Leon[8] has demonstrated that the food eaten by a mother rat influences the odours emanating from her caecal contents. His data further suggest that the pups were attracted to odours possessed by their mothers as a result of learning. Analogous hypotheses have been suggested by Hendry *et al.* for some insect pheromone systems[3]. In my experiments, the male guinea pig subjects experienced throughout their lives urine only from animals who had eaten guinea pig food. It is possible, therefore, that the preference shown for urine of guinea pigs fed GPF compared with urine of guinea pigs fed RF or synthetic food is the result of some aspect of previous experience with either guinea pig food or with urine from animals eating such food.

These data indicate that, in nature, male guinea pigs could discriminate among conspecific animals on the basis of previous dietary history. Although it is not surprising that dietary changes provide detectable differences in urinary chemical signals, the differential attractiveness of urine from animals on different diets shows that these differences have a communicative function. It is possible that information on the availability of a new food could be transmitted through urinary chemical signals. The importance of carefully controlling the diet in behavioural and biochemical studies of mammalian chemical communication is evident. Further, these data illustrate the complex nature of the information conveyed in guinea pig urine. In previous chemical studies, it has been found that separation of urine by a variety of procedures resulted in several fractions of widely different chemical character and molecular weight being responsible for its attractive nature of conspecific males[9]. It has been suggested[2] that a pattern of chemical substances may thus be involved in the urinary communication system of Cavia as well as in similar systems in other mammals. The data presented here further support these hypotheses.

Monell Chemical Senses Center,
University of Pennsylvania,
Philadelphia, Pennsylvania 19104

1. Birch, M. C. (ed.), *Pheromones* (North-Holland, Amsterdam, 1974).
2. Beauchamp, G. K., Doty, R. L., Moulton, D. G., and Mugford, R. A., in *Olfaction Reproductive Productive Processes and Behavior* (edit. by Doty, R. L.), (Academic, New York, in the press).
3. Hendry, L. B., Wichmann, J. K., Hindenlang, D. M., Mumma, R. O., and Anderson, M. E., *Science*, **188**, 59–62 (1975).
4. Rood, J. P., *Anim. Behav. Monog.*, **5**, 1–83 (1972).
5. Kunkel, P., and Kunkel, I., *Z. Tierpsychol.*, **21**, 602–641 (1964).
6. Beauchamp, G. K., *Physiol. Behav.*, **10**, 589–594 (1973).
7. Beauchamp, G. K., and Berüter, J., *Behav. Biol.*, **9**, 43–47 (1973).
8. Leon, M., *Physiol. Behav.*, **14**, 311–319 (1975).
9. Berüter, J., Beauchamp, G. K., and Muetterties, E. L., *Biochem. biophys. Res. Commun.*, **53**, 264–271 (1973).

Differentiation induced by cyclic AMP in undifferentiated cells of early chick embryo *in vitro*

THE undifferentiated cells of the chick embryo in the post-nodal piece (PNP), obtained by a transverse cut at 0.5 mm posterior to Hensen's node of stage 4 blastoderm, remain incapable of differentiation when grown in a variety of media or by chorioallontoic grafting technique. When combined with the region anterior to the streak, including Hensen's node, however, the PNP would undergo normal differentiation similar to the embryonic axial structures[1-3], suggesting that information might have been transferred from one tissue to another. Later studies by Butros[2] and Niu and coworkers[4,5] have shown that the PNP explants could be induced to develop into specific tissues, such as pulsating cardiac muscle tissues, when the explants were cultured in the presence of RNA isolated from the embryonic heart. The development was organ specific since RNA from kidney or thymus was incapable of inducing the heart formation[5]. To confirm this mode of differentiation in the PNP and to provide a meaningful biochemical interpretation to the phenomenon, we have begun studies examining the precise nature of the competent RNA and the sequence of events that leads to the formation of highly organised myofibrils. In a parallel study, we asked whether addition of exogenous cyclic AMP, which seems to be involved in control of morphological differentiation and biochemical changes in both normal and neoplastic cells (for review see ref. 6 and subsequent papers 7–15) would affect the development of the PNP in culture. Here we report that cyclic AMP, in the absence of exogenous RNA, can indeed produce specific morphological transformations, including the formation of heart-like pulsating tissues.

Fertile Leghorn chicken eggs from Spring Lake Farms, New Jersey, were incubated at 38 °C to obtain the definitive streak stage, stage 4 (ref. 16). The area opacua was trimmed off from the explanted blastoderms of the egg and the PNP was obtained by transecting the area pellucida at 0.6 mm behind Hensen's node to exclude totally the presumptive heart-forming region[17,18]. The PNPs collected from 20–30 eggs per experiment, were pooled together, and then divided randomly in batches of 3 each for *in vitro* cultivation. When the explants were cultured in the absence of cyclic AMP (control) the growth pattern of PNP was similar to that described previously[5,19]. The primitive streak disappeared after 24 h of incubation and patches of red blood cells appeared on the third day. No axial structures and twitching tissues were seen in any of the explants even after 10 d of incubation (Fig. 1*a*). The absence of twitching tissues in all the control explants (see Table 1) indicated
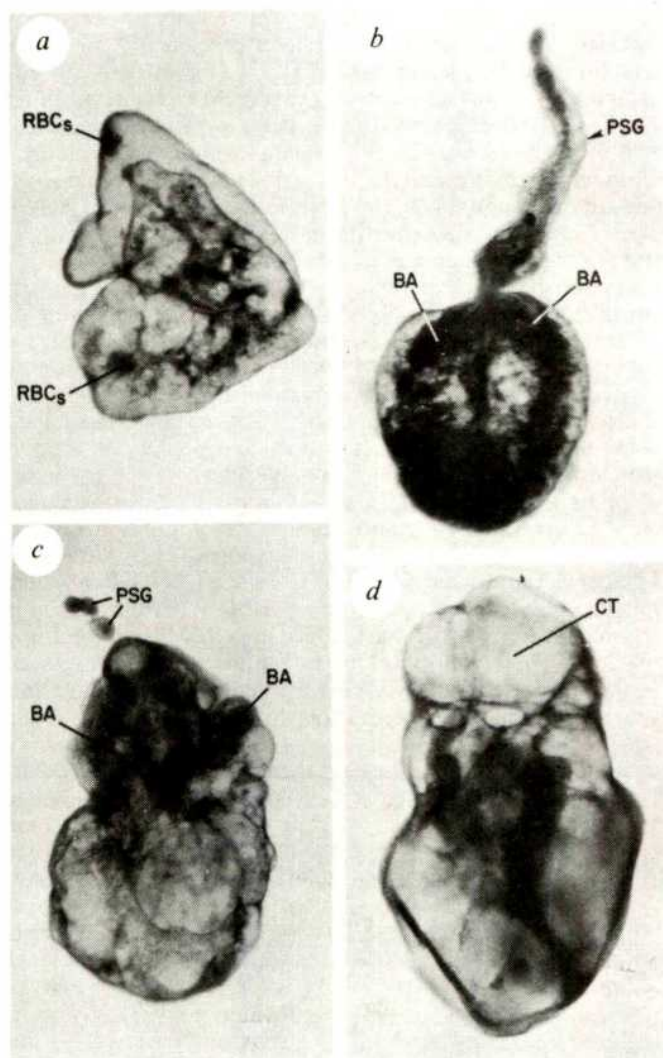
Fig. 1 Whole mounts of PNP cultured with or without cyclic AMP (0.5 mM). The cultivation was a modification of New's technique[31] as described earlier[32]. Three PNPs were placed on the yolk-free vitelline membrane (VM) previously mounted around a glass ring, with the epiblast in contact with the VM. After removing excess PC (Pannett-Compton) solution[32] the PNPs were carefully flattened on the VM. One millilitre of the nutritive medium (NM)[19] was then pipetted in each dish containing 3 PNP. The medium contained 50 U ml^{-1} of penicillin and 50 µg ml^{-1} of streptomycin. The pH of the final solution after adding the supplements was 7.6. Cyclic AMP (0.5 mM) was added directly to the medium. Incubation was done at 38 °C and the PNPs were examined at regular intervals. All operations were carried out in sterile conditions. The PNPs were examined at regular intervals with a binocular microscope for changes in growth, appearance of the pulsating tissues and general morphology. The PNPs were photographed after staining with haemotoxylin and eosin Y at times indicated. a, −cyclic AMP, 72 h incubation; b, + cyclic AMP 40 h incubation; c, + cyclic AMP, 72 h incubation; d, +cyclic AMP, 96 h incubation. RBCs, patches of red blood cells; PSG, primitive streak outgrowth; BA, beating area; CT, cardiac tube. Magnification ×30.

that transection at 0.6 mm posterior to the node did exclude the entire presumptive heart-forming region from the PNP. In contrast, the addition of cyclic AMP (0.5 mM) caused striking changes in growth and development of the PNP. Not only did the primitive streak persist during the early stages, but it eventually grew to be equal in length to the whole explant, and in addition, a median structural outgrowth (PSG) projecting from the anterior end of the streak became visible during the first 24 h of incubation. The site

of the outgrowth projection, that is, the anterior end of the primitive streak, was similar for all the cyclic AMP-induced PNPs. Further incubation resulted initially in the appearance of infrequent twitchings near the basal portion of the outgrowth (Fig. 1b) which changed on day four, into rhythmic and spontaneous pulsations resembling that of cardiac tissues in normal intact embryos. This was accompanied by a progressive digression of the PSG, which was first transformed into a beaded string-like structure breaking off eventually from the rest of the explant (Fig. 1c). After 4 d of incubation cardiac tubular structures with rhythmic pulsations were visible (Fig. 1d). The steady rate of pulsations lasted 2–4 d without supplementing the medium. An important difference between the cyclic AMP-induced growth of the nodeless streak and that of normal intact embryo was the absence of normal regression of the primitive streak in the former.

An ultrastructural examination of the PNPs (Fig. 2) revealed that typical epithelial, mesenchymal and erythroblastic cells were found in explants of both control and experimental groups. Scantily-dispersed endoplasmic reticulum, vesicular nuclei and a few mitochondria appeared to be the characteristic features of the undifferentiated PNP. However, all randomly selected PNPs grown with effective concentrations of cyclic AMP (see below) contained, in addition to the normal components, highly differentiated myoblasts which seemed to be similar in ultrastructural complements to the normal chick embryonic myocardial cells. Clearly evident were the

Table 1 Effect of cyclic AMP and analogues on differentiation in PNP

Additions	No. of PNP treated	No. of differentiated PNP*
Cyclic AMP (0.05 mM)	12	0 (0%)
Cyclic AMP (0.1 mM)	12	1 (8%)
Cyclic AMP (0.5 mM)	39†	29 (74%)
Theophylline (0.5 mM)	6	0 (0%)
Theophylline (1.0 mM)	6	0 (0%)
Theophylline (2.0 mM)	10	6 (60%)
Cyclic AMP (0.1 mM) + theophylline (0.5 mM)	6	4 (66%)
Dibutyryl cyclic AMP (0.05 mM)	3	0 (0%)
Dibutyryl cyclic AMP (0.1 mM)	3	0 (0%)
Dibutyryl cyclic AMP (0.5 mM)	9	2 (22%)
Dibutyryl cyclic GMP (0.5 mM)	6	0 (0%)
Adenosine (0.1 mM)	6	0 (0%)
Adenosine (0.2 mM)	6	5 (83%)
Adenosine (0.5 mM)	19†	15 (79%)
Adenosine (0.2 mM) + theophylline (1.0 mM)	6	0 (0%)
Adenosine (0.5 mM) + theophylline (1.0 mM)	6	4 (66%)
No additions (control)	84†	0 (0%)

Details of culture condition were described in the legend to Fig. 1.

*Differentiation was characterised by the appearance of spontaneous and rhythmic pulsations after 96 h incubation. In addition, characteristic myofibrillar structures were found in all randomly-selected pulsating explants when examined in the electron microscope. None of the explants of control series or those treated with non-effective agents contained myofibrils.

†The number of PNP explants represents a cumulative total of several experiments.

orderly arranged myofibrils with intercellular apical junctions or intercalated disks as described previously[5,20]. These myofibrils were absent in the PNPs of control group and in those which did not differentiate into beating tissues. Measurements of the total protein content[21] using three
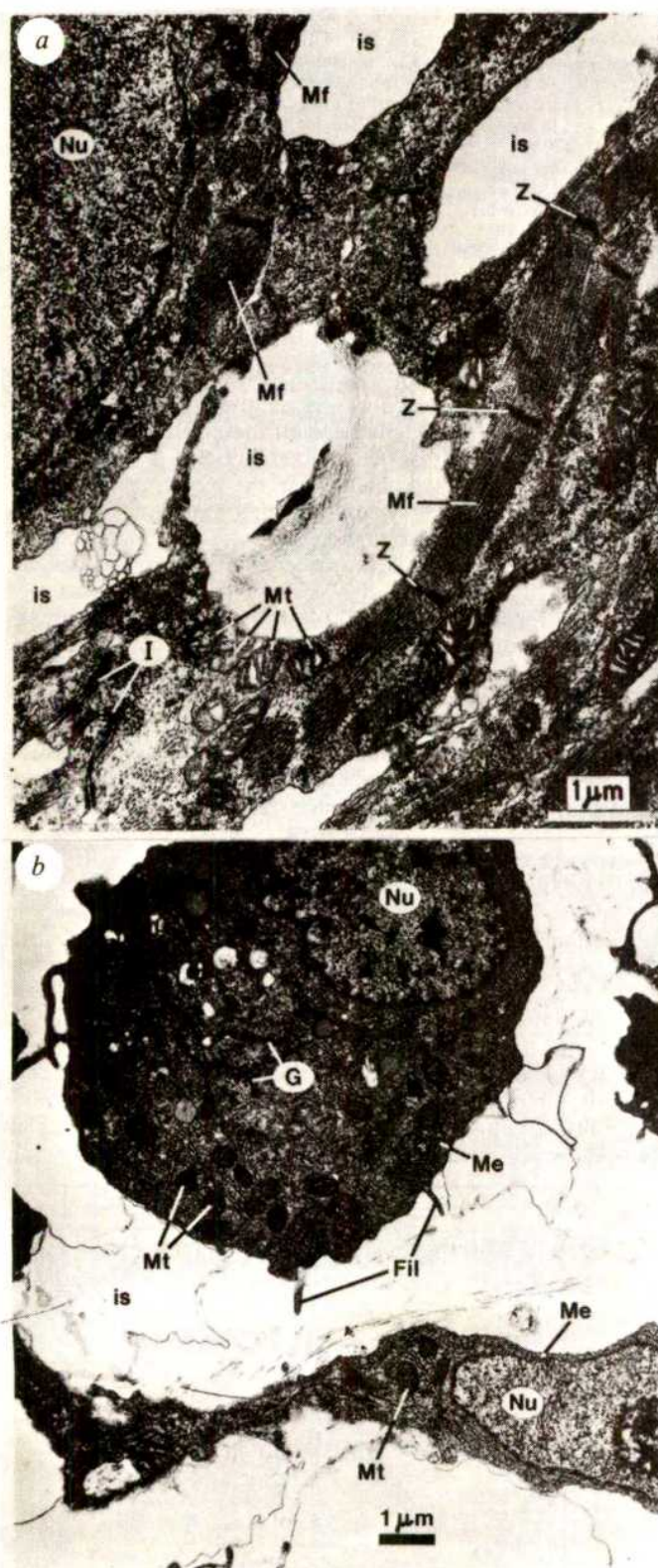
Fig. 2 Electron micrographs of cells of PNP after 96 h incubation with or without cyclic AMP. Electron microscopy was done in Dr M. Boublik's laboratory at the Roche Institute of Molecular Biology according to O'Brien *et al.*[34]. *a*, +cyclic AMP (0.5 mM); *b*, −cyclic AMP; Nu, nucleus; MF, myofibrils; Z, Z-bands; is, intercellular space; I, intercalated disk; Mt, mitochondria; G, Golgi body; Fi, filopodia; Me, mesenchyme.

PNP explants per assay, done in triplicate, revealed that each PNP of stage 4 blastoderm contained $9.0 \pm 1.0 \mu g$ protein, which increased to $21. \pm 0.7 \mu g$ after 4 d of incubation in the absence of cyclic AMP. However, the PNP which differentiated into beating tissues contained $30.6 \pm 1.3 \mu g$ per protein per explant. Thus, the formation of myofibrillar structures, the acquisition of spontaneous and rhythmic pulsations, the appearance of distinct morphological changes, such as the persistence of the primitive streak, and the increase in total protein content were all characteristic features of only those PNPs which received effective amounts of cyclic AMP and provided clear lines of evidence for a differentiative transition in the PNP.

Cyclic AMP, at 0.5 mM, was most effective, whereas concentrations below this level caused little or no differentiation (Table 1). Dibutyryl cyclic AMP, on the other hand, when used at similar concentrations appeared to be a poor inducer. Theophylline, which acts specifically as an inhibitor of phosphodiesterase and thereby increases the intracellular level of cyclic AMP[22], also induced differentiation similar in microscopical examination to that caused by cyclic AMP. Furthermore, a synergistic effect was clearly evident when non-effective concentrations of cyclic AMP (0.1 mM) and theophylline (0.5 mM) were added together to the culture medium. These results led us to speculate that changes in intracellular levels of cyclic AMP might be responsible for the specific morphogenetic transformations in the PNP.

Table 2 Endogenous levels of cyclic AMP in cultured PNP explants

Incubation Time (h)	Experiment	No additions	Cyclic AMP (Pmol μg⁻¹ protein) +cyclic AMP	+ADO	+THEO
0	1	0.019	0.023	0.020	0.026
	2	0.022	0.030	–	–
12	1	0.030	0.490	0.070	0.100
	2	–	0.460	0.066	0.088
48	1	0.058	0.730	0.040	–
	2	0.044	0.520	–	–
96	1	0.050	0.690	0.125	0.091

Endogenous levels of cyclic AMP were essentially determined according to Gilman[35] and Brown *et al.*[36] and as described in Amersham/Searle *Cyclic AMP Assay Kit*, TRK 432. Exogenous cyclic AMP, Ado (adenosine), or Theo (Theophylline), 0.5 mM each, were added directly to the culture medium (NM) and the PNPs were placed on top of the vitelline membrane (VM) as described in the legend to Fig. 1. At times indicated the explants were removed, washed 4 times with ice-cold Ringer's solution. Removal of extracellular cyclic AMP was indicated by no further decrease in cyclic AMP with additional washes. The PNPs were homogenised in 5% trichloroacetic acid. The precipitate recovered after centrifugation was used for protein determination[21] and the supernatant was extracted thoroughly with ether and dried by lyophilisation. The dried material was taken up in Tris-EDTA buffer for cyclic AMP determination. Three PNPs were used per assay and the assays were done in duplicate for each experiment. The numbers above represent an average of six PNPs for each time point.

This assumption was tested by (1) a direct estimation of endogenous cyclic AMP levels and (2) by the exogenous addition of adenosine which is known to cause an elevation of cyclic AMP levels in a variety of tissues[23-25], because of the direct activation of the enzyme adenylate cyclase by adenosine[26]. Intracellular cyclic AMP level was estimated as described in the legend to Table 2. The explants after incubation with exogenous cyclic AMP were washed 4 times with ice-cold Ringer's solution. Removal of all extracellular cyclic AMP was indicated by no further decrease in cyclic AMP with additional washes. When cyclic AMP was added to the culture medium, there was a 25-fold increase in the intracellular cyclic AMP level after 12 h incubation (Table 2). This increase represented about 70% of the total increase observed at 48 h incubation, which was the saturation level of cyclic AMP uptake. The 12-h-old explants were almost one-third in size and in protein content compared with those

cultured for 96 h. It would seem therefore that the uptake of cyclic AMP occurred predominantly during the early stages of incubation and was not directly related to the growth and increase in size of the explants

Addition of theophylline and adenosine to the culture medium also caused a significant increase in the intracellular cyclic AMP levels (Table 2). The increase was, however, substantially lower than that observed with exogenous cyclic AMP, but was apparently sufficient to cause the differentiative transition in the PNP (see Table 1). A specific feature of the adenosine effect is the fact that it can be abolished by theophylline (methylxanthine), which acts as a competitive inhibitor of adenosine[24,26,27]. As shown in Table 1, adenosine, but not guanosine, could mimic the effect of cyclic AMP. The concentration of adenosine required for an optimum effect was lower than that required for cyclic AMP. This might be due to a differential uptake of these compounds by the PNP in the culture conditions used. The adenosine-stimulated differentiation was blocked effectively by theophylline. The inhibition caused by theophylline was diminished when adenosine concentration was raised from 0.2 to 0.5 mM, indicating the competitive nature of these drugs. These effects of theophylline and adenosine on PNP were strikingly similar with those observed on cyclic AMP system in other tissues. Thus, the results taken together strongly suggested that the observed changes in the PNP growth and development were caused by modulations in intracellular levels of cyclic AMP.

The multiplicity of the effects caused by cyclic AMP was reflected by the formation of highly organised myofibrillar structures, acquisition of spontaneous pulsations and by the appearance of specific morphological changes described above. A recent report[30] has confirmed these findings and has shown that in addition to heart-like tissues, neural tissues, notochord and nephric tubules were also noted. The concentrations of cyclic nucleotides required, however, were 6–30-fold higher than those used in our studies. Reporter and Rosenquist[29] have previously reported that cyclic AMP levels in the precardiac mesodermal area in head-process stage of the chick embryo was about 20 times higher than in kidney-forming, liver mesoderm, and in neuro-epithelial ectoderm regions. The authors suggested that the regional differences in cyclic AMP concentrations may indicate the possible involvement of cyclic AMP, a presumed morphogen, in early chick embryonic development. We believe that there is a relationship between the state of development of the cultured PNP explants and the changes in cyclic AMP levels together with the cyclic AMP-associated properties of cells, such as the changes in membrane permeability assembly of microtubules, and so on. But the relationship between the exogenous RNA-induced mode of differentiation[3,4,5,36] and that caused by cyclic AMP is yet to be ascertained.

We thank Dr M. Boublik and Mr F. Jenkins for doing the electron-microscopic studies and Dr A. Blume for many helpful discussions. We also thank Mrs Dallas George for secretarial assistance.

A. K. DESHPANDE
M. A. Q. SIDDIQUI

*Department of Biochemistry,
Roche Institute of Molecular Biology,
Nutley, New Jersey 07110*

1 Rawles, M. E., *Physiol Zool.*, 16, 22–44 (1943)
2 Butros, J., *J. exp. Zool.*, 149, 1–20 (1962).
3 Butros, J., *J. Embryol exp. Morph.*, 13, 119–128 (1965)
4 Sanyal, S., and Niu, M. C., *Proc. natn Acad. Sci U S A.*, 55, 743–750 (1966).
5 Deshpande, A. K., Niu, L. C., and Niu, M. C., in *The Role of RNA in Reproduction and Development* (edit by Niu, M. C., and Segal, S.), 229–246 (Elsevier, New York, 1973).
6 Pattan, L., and Johnson, G. S., in *Advances in Cancer Research*, 19 (edit. by Klein, G., Weinhouse, S., and Haddow, A.), 303–329 (Academic Press, New York, 1974).

7 Rosen, F. J., Murray, R. A., Pichichero, M. E., and Braden, W. G., *Science*, 175, 73–74 (1972)
8 Wahrmann, J. P., Winland, R., and Luzzati, D., *Nature new Biol.*, 245, 112–113 (1973)
9 Hovi, T., and Vaheri, A., *Nature new Biol.*, 245, 175–176 (1973).
10 Seifert, W. E., and Rudland, P. S., *Nature*, 248, 138–139 (1974).
11 Rudland, P. S., Gospodarowicz, D, and Seifert, W, *Nature*, 250, 741–742 (1974).
12 Weinstein, Y., Chambers, D. A, Bourne, H. R., and Melmon, K. L., *Nature*, 251, 352–35 (1974).
13 Bjerre, B., *Experientia*, 30, 534–535 (1974).
14 Filosa, S, Pictet, R, and Rutter, W. J., *Nature*, 257, 702–704 (1975).
15 Bossing-Schneider, R, *Nature*, 256, 137–138 (1975).
16 Hamburger, V, and Hamilton, H. L., *J. Morph.*, 88, 49–92 (1951).
17 Rudnick, D., *Anat. Rec.*, 70, 351–368 (1938).
18 Mulherkar, L, *J. Embryol exp. Morph*, 6, 1–14 (1958).
19 Niu, M C., and Deshpande, A. K, *J Embryol exp. Morph.*, 29, 485–501 (1973).
20 Manasek, F J., *J. Morph.*, 125, 329–366 (1968).
21 Lowry, O H, Rosebrough, N. J., Farr, A. L., and Randall, R. J., *J. biol Chem.*, 193, 265–275.
22 Butcher, R W, and Sutherland, E W., *J biol Chem*, 237, 1244–1250 (1962).
23 Schultz, J., and Daly, J W., *J biol. Chem*, 248, 853–859 (1973).
24 Blume, A. J., Dalton, C., and Sheppard, H., *Proc. natn. Acad. Sci. U.S.A.*, 70, 3099–3102 (1973).
25 Clark, R B, Gross, R., Ying-fu, Su, and Perkins, J. P., *J. biol. Chem.*, 249, 5296–5303 (1974).
26 Blume, A J., and Foster, C J., *J. biol Chem.*, 250, 5003–5008 (1975).
27 Sattin, A., and Rall, T. W, *Molec Pharmac*, 6, 13–23 (1970).
28 Lee, H, Kalmus, G. W., and Nagele, R. G, Jr, *In Vitro*, 21, 286 (abstract) (1976).
29 Reporter, M, and Rosenquist, G. C., *Science*, 178, 628–630 (1974).
30 Deshpande, A K, and Siddiqui, M A. Q., *Devl Biol.* (in the press)
31 New, D A. T., *J Embryol. exp Morph*, 3, 326–331 (1955).
32 Chauhan, S P. S., and Rao, K. V., *J Embryol exp. Morph*, 23, 71–78 (1970).
33 Pannett, C A., and Compton, A, *Lancet*, 206, 381–384 (1924).
34 O'Brien, R. A, Boublik, M, and Spector, S. J., *Pharmac. exp. Ther.*, 194, 145–153 (1975).
35 Gilman, A. G., *Proc. natn. Acad. Sci U.S A.*, 67, 305–312 (1970).
36 Brown, B. L., Albano, J. D. M., Ekins, R. P., and Sgherzi, A. M, *Biochem. J.*, 121, 561–562 (1971).

Postnatal development of the synaptic organisation of the lateral geniculate nucleus in the kitten with unilateral eyelid closure

CLOSURE of the eyelids over one eye during a well defined period of susceptibility in kittens has been shown to result in severe structural changes of the nerve cells in the lateral geniculate nucleus[1,2]. Neurones in those parts of the lateral
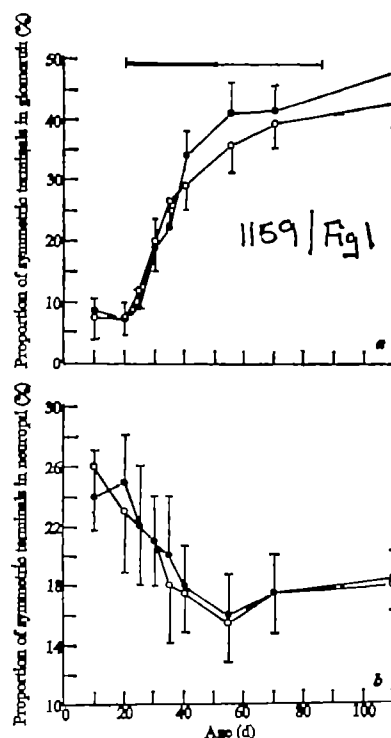
Fig. 1 The change in proportion of the symmetric synapses with age in the glomeruli (a) and the extraglomerular neuropil (b) of the binocular parts of lamina A and A1 of the lateral geniculate nucleus in the normal kitten. The phase of rapid change occurs during the period of maximum susceptibility. Vertical bars indicate standard error; ●, lamina A, ○, lamina A1; thick part of the horizontal bar indicates period of maximum susceptibility and thin part the period of declining sensitivity.

geniculate nucleus which are related to the binocular visual field and which receive fibres from the closed eye are smaller than those in the laminae innervated from the open eye. The fact that cellular changes do not occur in the part of the nucleus related to the monocular visual field has led to the hypothesis that there is a process of competition or an imbalance between the neurones in the deprived and stimulated portions of the binocular parts of the laminae, whereas such competition is not present in the monocular segment[2,3]. We have examined the lateral geniculate nucleus with the electron microscope at different intervals extending through the period of susceptibility after monocular eyelid closure and in a parallel series of normally reared kittens.

In nine kittens the eyelids over one eye were sutured together at the time of eye opening at about 10 d of age, and these animals were allowed to survive for varying periods, the longest survival period being 100 d. These animals were perfused at 24 °C with a mixture of 4% paraformaldehyde and 1% glutaraldehyde after a brief washout with a balanced salt solution. The lateral geniculate nucleus of both sides was removed from the thalamus, cut coronally into three blocks and processed for electron microscopy. After examination of 1-μm thick sections stained with toluidine blue[4], ultrathin sections were cut including the optic radiation and adjoining parts of laminae A and A1 and from the monocular segment of lamina A. In all animals the sections were taken from the middle third of the anteroposterior extent of the nucleus, from about the middle of the mediolateral extent of laminae A and A1 and from the monocular segment[2] in the lateral part of lamina A.

A study of the monocularly deprived lateral geniculate nucleus with the electron microscope showed that the de-prived laminae underwent little qualitative change compared with the undeprived laminae, and as the differences in cell sizes are not so opparent in ultra-thin sections it was often not possible to determine which were the deprived and undeprived laminae. There was no change in the appearance of the cell somata and the constituent organelles, apart from a slight diminution in the amount of the granular endoplasmic reticulum, in the dendritic profiles, axons, axon terminals and glia; in particular the large retinal terminals showed no abnormalities either in their electron density or synaptic vesicle content. Thus, there is no evidence for degeneration either of the axons from the deprived retina, or of the constituent neurones of the nucleus, nor was there a glial reaction. In view of this unexpected absence of qualitative changes a quantitative study was made, and because of the evidence suggesting that the cells of the undeprived laminae in such experimental animals may have undergone some hypertrophy[3,5,6], it was considered necessary to compare the quantitative data in this series with corresponding data from the laminae in normal kittens of the same ages. The brains of nine normal kittens were therefore perfused in the same manner and ultra-thin sections for electron microscopy taken from corresponding parts of the lateral geniculate nucleus.

Within the glomeruli the numbers and proportions of the asymmetric and symmetric synapses were counted, and their postsynaptic profiles were recorded at each age. No attempt was made to differentiate between the symmetric contacts formed by axon terminals and presynaptic dendrites. About 25 glomeruli were sampled randomly from the binocular parts of each of the laminae of the lateral geniculate nucleus of both sides in the experimental brains, and of one side of the normal brains, and similar data from the monocular segments of lamina A were obtained. Between
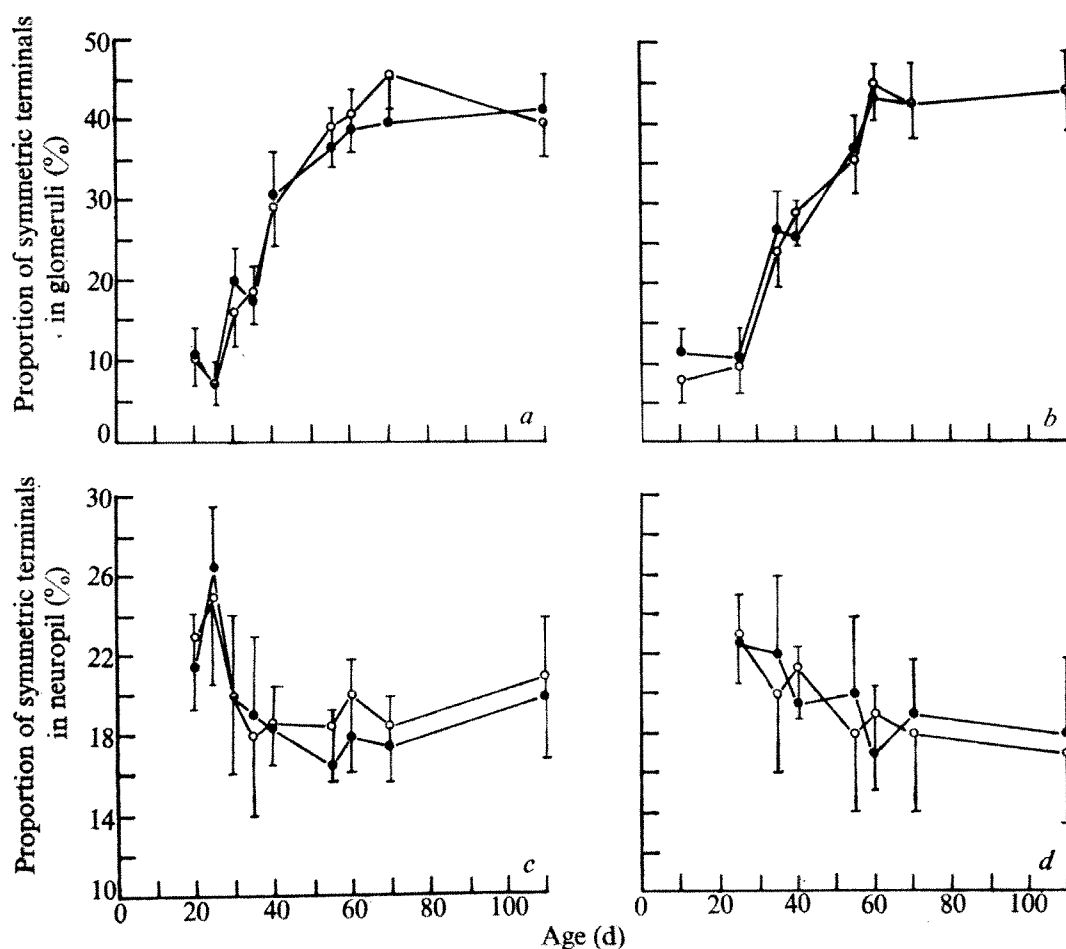
Fig. 2 The change in proportion of symmetric synapses with age in the glomeruli (*a* and *b*) and in the extraglomerular neuropil (*c* and *d*) of the binocular segments of laminae A and A1 of the ipsilateral (*a* and *c*) and contralateral (*b* and *d*) lateral geniculate nucleus after monocular eyelid closure. On the ipsilateral side lamina A1 has been deprived, and on the contralateral side lamina A. ●, lamina A; ○, lamina A1.

100 and 150 synapses were classified into these categories in each lamina, the total sample amounting to approximately 10,000 synapses. These counts were made from micrographs at a magnification of 16,000. The proportions of these two types of synapses were also estimated in the extraglomerular neuropil, approximately the same number being sampled as in the glomeruli; they were classified at the time of study on the microscope.

Within the glomeruli of the normal kittens there is a phase of rapid increase in the proportion of the symmetric synapses from less than 10% at 20 d to between 35 and 40% at 55 d, with relatively little further change up to 110 d (Fig. 1a) In contrast, in the extraglomerular neuropil, there is a steep fall in the proportion of symmetric synapses from 25 to 15% between 20 and 55 d (Fig. 1b). These figures confirm the qualitative observations that between 20 and 55 d there is a differentiation of the profiles within glomeruli, especially an increase in the number of vesicle-filled presynaptic dendrites making symmetrical synapses; during this period the glomeruli also became more clearly segregated by their glial envelopes. The time course and degree of maturation finally attained is similar in both main laminae. The comparable data from the glomeruli and the neuropil from the experimental brains show no significant (two-tailed *t* test, *P*>0.05) differences in the rate of development between the laminae in the normal and experimental brains and between the deprived and undeprived laminae on the two sides of the same experimental brains (Fig. 2). The

Fig. 3 A comparison of the time course of the change in proportion of symmetric synapses (●) with the other morphological changes which occur in lamina A of the lateral geniculate nucleus and optic tract during postnatal development in the kitten. Cell size (□) from Garey *et al.*[8], synaptic density (■) from Cragg[10] and myelination of optic tract fibres (○) from Moore *et al.*[11]. Horizontal bar indicates period of susceptibility.

proportions of the two types of synapses within the glomeruli and the neuropil of the monocular segment follow the same course of development as the binocular segments and there are probably no significant differences (*P*>0.05) in the deprived and normal monocular segments.

An analysis has been made of the types of profiles post-synaptic to the asymmetric and symmetric synapses respectively within the glomeruli. At younger ages there are more relay cell dendrites receiving asymmetric synapses than presynaptic dendrites (for example, at 20 d 90% of the profiles postsynaptic to the asymmetric retinal boutons are relay cell dendrites), but in progressively older animals there is an increasing number of presynaptic dendrites receiving such synapses (at 70 d about 70% of the postsynaptic profiles are relay cell dendrites). This difference, however, may be more apparent than real because of the greater difficulty

in distinguishing between relay and presynaptic dendrites at earlier ages when the presynaptic dendrite is not so clearly differentiated. Similar comparisons cannot be made about the profiles postsynaptic to symmetric synapses because the number of presynaptic dendrites making synapses at younger ages is very small, and they are less definitely identifiable, but at 110 d the majority of the symmetric synapses, like the asymmetric synapses, are on the relay cell dendrites. No statistically significant differences (*P*>0.05) were found in the postsynaptic profiles between the monocular and binocular segments, nor between the laminae of normal brains and the undeprived and deprived laminae of the brains of the experimental animals of the same ages.

The numerical data used for the estimates of the proportions of the types of synapses and of the postsynaptic profiles have also been used to provide estimates of the mean numbers of synapses per glomerulus at different ages. These data confirm the qualitative impression that the number of symmetric synapses in a glomerulus increases with age (for example the mean number at 20 d is 0.33 (s.e. 0.47), whereas at 110 d it is 2.76 (s.e. 1.76) in lamina A; *P*<0.001). The mean number of retinal synapses does not show a significant change with age, so the explanation for the change in the proportions of the synapses that has already been noted is due mainly, and perhaps solely, to an increase in the number of symmetric synapses. The brains of the experimental series showed similar changes with age and significant differences (*P*>0.05) were not found between them and those of the normal brains.

These electron microscopic observations show that there is a marked degree of development of the synaptic organisation occurring during the period of susceptibility to effects of eyelid closure, but that there are no appreciable differences to be found between the nucleus in a normal brain and in those animals subjected to monocular eyelid closure. Thus, if the shrinkage of the neurones in the lateral geniculate nucleus seen with the light microscope is associated with alterations in the retinal terminals or in the synaptic organisation, it must be at a level more subtle than that studied here. The cell shrinkage may be related to features of the intrinsic organisation, such as reciprocal and serial synapses, or to changes in the number of synaptic vesicles within terminals. It may, on the other hand, not be related to change in the lateral geniculate nucleus, but at the site of termination of the geniculo-cortical fibres in the visual cortex[7,8]. The findings in the series of normal brains of a marked change in the proportions of synapses within glomeruli and the extraglomerular neuropil between the ages of 20 and 55 d is further evidence that the nucleus is undergoing a phase of rapid morphological development during the period of susceptibility to monocular deprivation. It has already been shown that during this period there is a rapid growth in the size of cell bodies within the nucleus[9], in the synaptic density[10], and in the degree of myelination of the incoming optic tract fibres[11] (Fig. 3). It is of interest that within the glomeruli the change in proportion is due to an increase in the number of symmetric terminals, which, on present evidence, arise from interneurones[12,13], whereas in the neuropil the asymmetric synapses increase relatively more and these are probably predominantly the terminals of cortico-geniculate axons and collaterals of geniculo-cortical relay cells[14,15]. A change, however in the proportion of intrinsic axons after monocular and binocular eyelid suture has been found in another major site of termination of retinal fibres, the superior colliculus in the rat, where the final phase of growth of intrinsic axons fails to occur after eyelid closure[16]. This presents a paradox that within the superior colliculus unilateral eyelid closure has an effect on the attainment of the mature synaptic organisation, although no appreciable shrinkage of cells has been found, yet in the lateral geniculate nucleus, where there is a marked shrinkage of the neurones, there does not seem

to be any accompanying change in the synaptic organisation.

D. A. WINFIELD
M. P. HEADON
T. P. S. POWELL

Department of Human Anatomy,
University of Oxford,
South Parks Road,
Oxford OX1 3QX, UK

1 Wiesel, T. N., and Hubel, D. H., *J. Neurophysiol.*, **26**, 978–993 (1963).
2 Guillery, R. W., and Stelzner, D. J., *J. comp. Neurol.*, **139**, 413–422 (1970).
3 Guillery, R. W., *J. comp. Neurol.*, **144**, 117–130 (1972).
4 Richardson, K. C., Jarrett, L., and Finke, L. H., *Stain. Tech.*, **35**, 313–323 (1960).
5 Garey, L. J., Fisken, R. A., and Powell, T. P. S., *Brain Res.*, **52**, 363–369 (1973).
6 Sherman, S. M., and Widon, J. R., *J. comp. Neurol.*, **161**, 183–196 (1975).
7 Garey, L. J., and Pettigrew, J. D., *Brain Res.*, **66**, 165–172 (1974).
8 Vrensen, G., and De Groot, D., *Brain Res.*, **93**, 15–24 (1975).
9 Garey, L. J., Fisken, R. A., and Powell, T. P. S., *Brain Res.*, **52**, 359–362 (1973).
10 Cragg, B. G., *J. comp. Neurol.*, **160**, 147–166 (1975).
11 Moore, C. L., Kalil, R., and Richards, W., *J. comp. Nerol.*, **165**, 125–136 (1976).
12 Famiglietti, E. V., and Peters, A., *J. comp. Neurol.*, **144**, 285–334 (1972).
13 Lieberman, A. R., *Brain Res.*, **59**, 35–59 (1973).
14 Guillery, R. W., *Z. Zellforsch.*, **96**, 1–38 (1969).
15 Jones, E. G., and Powell, T. P. S., *Proc. R. Soc.*, **B172**, 173–185 (1969).
16 Lund, J. S., and Lund, R. D., *Brain Res.*, **42**, 21–32 (1972).

Partial replacement of serum by selenite, transferrin, albumin and lecithin in haemopoietic cell cultures

CULTURE media for mammalian cells usually require supplementation with serum to supply as yet undefined needs. Because the requirements are likely to be multiple, it is difficult to distinguish biologically between the effects of nonspecific 'nutritional' factors in serum and those of specific regulatory factors. Partial replacement of serum in these systems by chemically defined substances supplying the nonspecific needs would therefore represent a significant advance towards the definition of such specific factors. We have examined the role of several serum components in cultures of freshly explanted haemopoietic cells with this goal in mind.

Red-cell precursors in freshly explanted mammalian bone marrow will proliferate to form colonies if the medium contains serum and the glycoprotein hormone erythropoietin[1,2]. Similarly, colony formation by granulocyte/macrophage precursors is dependent on serum as well as a specific glycoprotein colony-stimulating factor[3]. We decreased the concentration of serum in these cultures until growth was limited. The serum concentration was made the only variable by maintaining erythropoietin or colony-stimulating factor at high and non-limiting levels. A mixture of a large number of known serum constituents was added to the cultures and restored growth. Elimination of the components one by one then established which were active and which were not. Four components—sodium selenite, transferrin, bovine serum albumin (BSA) and lecithin—accounted for all the activity of the mixture. We demonstrate here that in combination these substances replace most of the serum required for granulocyte/macrophage and erythroid colony growth and also facilitate some granulocyte/macrophage proliferation without added serum.

Bone marrow cells from the femurs of BDF₁ mice were plated in modified Dulbecco's medium (legend Fig. 1) made viscous with methyl cellulose[2]. Erythroid colonies of eight or more cells were identified[2] after 2 d of incubation, and granulocyte/macrophage colonies of more than 100 cells were scored after 8 d.

In these conditions, maximum numbers of erythroid colonies required the addition of 30% foetal calf serum. When the serum concentration was decreased to 1%, no colonies formed (Table 1). At this serum concentration, selenite, BSA and transferrin added individually had little

effect, while in combination they stimulated the formation of large numbers of colonies. Addition of lecithin to this combination caused no further enhancement. Omission of either selenite, BSA or transferrin from the mixture resulted in a significant decrease in colony growth, while omission of serum almost completely abolished growth.

Maximum numbers of granulocyte/macrophage colonies ordinarily required the addition of 10–15% serum. When
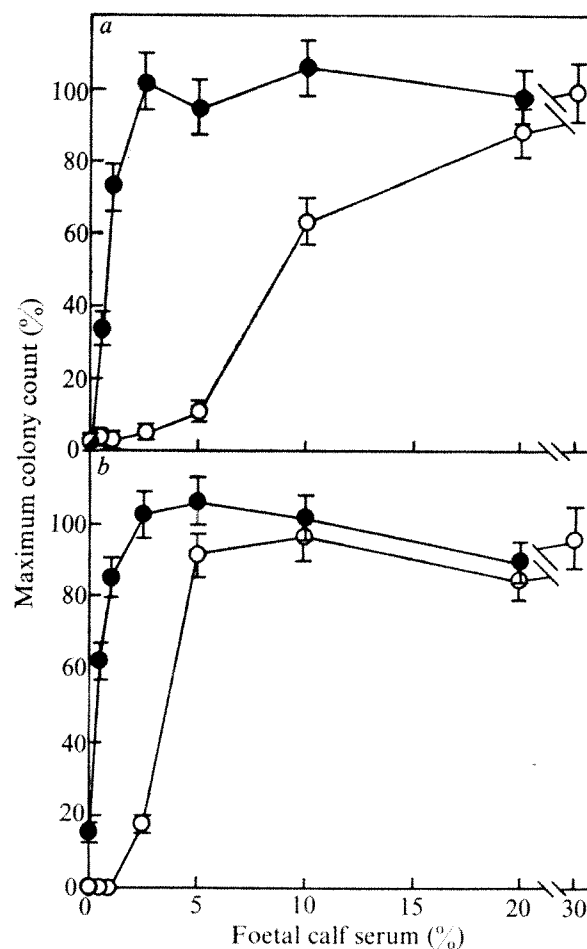
Fig. 1 Dependence of colony formation on foetal calf serum concentration. *a*, Erythroid colonies, percentage of counts at 30% serum (240 per 10⁵ cells), 0.5 erythropoietin units per ml. ○, Nothing added; ●, supplemented with 0.75% BSA, 3.4×10^{-6} M transferrin $+ 1.6 \times 10^{-6}$ M FeCl₃, and 10^{-7} M Na₂SeO₃. *b*, Granulocyte/macrophage colonies, percentage of counts at 30% serum (126 per 10⁵ cells), plateau amounts of colony stimulating factor. ○, Nothing added; ●, supplemented with BSA, transferrin/FeCl₃, Na₂SeO₃ and 3×10^{-5} M egg lecithin. Bars indicate standard errors. Cells were cultured at 10⁵ per ml in Dulbecco's modified Eagle's MEM (Gibco H21) containing 0.8% methyl cellulose and supplemented with 10^{-4} M α-thioglycerol plus the following (µg ml⁻¹): L-alanine (25), L-asparagine.H₂O (50), L-aspartic acid (30), L-cysteine (70), L-glutamic acid (75), L-proline (40), Na pyruvate (110), vitamin B₁₂ (0.025) and biotin (0.03). Human urinary erythropoietin was adsorbed on to benzoic acid and further purified on DEAE cellulose[4] and Sephadex G-100 (ref. 2). Colony-stimulating factor from serum-free mouse kidney cell conditioned medium[5] was partially purified by two passages over Sephadex G-150. BSA (Behringwerke, electrophoretically pure) was deionised over a mixed bed ion-exchange resin[6], charcoal (Norit A) extracted[7] (pH 3, 55 °C, 30 min) and chromatographed on an AcA 34 gel column to isolate the monomeric form. Human transferrin (Behringwerke) was further purified on AcA 34 and prepared as a 125× stock (38 mg lyophilised transferrin $+ 48$ µg FeCl₃.6H₂O per ml) in Dulbecco's medium. Na₂SeO₃ (Merck) was prepared as a 125× stock solution (2.2 µg ml⁻¹) in Dulbecco's medium containing 5% BSA. Egg lecithin (BDH, > 95%) was dissolved in propylene glycol (25 mg ml⁻¹) and diluted 10-fold to a 125× stock solution in Dulbecco's medium containing 5% BSA. Plate concentrations of propylene glycol (0.08%) were not inhibitory. Foetal calf serum was used without treatment.

Table 1 Effect of BSA, transferrin (Tf) and selenite (Se) on erythroid colony counts

Addition	Colonies per 10⁵ nucleated cells
30% Foetal calf serum (FCS)	353 ± 13*
1% FCS	0
1% FCS, Se	6 ± 3
1% FCS, BSA	13 ± 5
1% FCS, Tf	4 ± 2
1% FCS, Se, BSA, Tf	232 ± 9
— , Se, BSA, Tf	7 ± 3
1% FCS, —, BSA, Tf	117 ± 6
1% FCS, Se, — , Tf	42 ± 7
1% FCS, Se, BSA, —	67 ± 10

Concentrations of BSA, Tf and Se were as in Fig 1.
*Standard error.

serum was eliminated, no colonies formed (Table 2). Selenite, BSA, transferrin or lecithin had no effect when added individually to the serum-free medium. When added in combination, however, they stimulated growth of significant numbers of colonies. Although the colonies obtained in these conditions were generally smaller than those obtained with 15% serum, many contained at least 1,000 cells and were visible to the unaided eye. Omission of either BSA or transferrin from the mixture eliminated colony formation entirely and colony numbers were also reproducibly lower in the absence of selenite or lecithin.

The serum-sparing effect of these substances is demonstrated in Fig. 1. The addition of selenite, BSA and transferrin reduced the serum required for maximum numbers of erythroid colonies from 30 to 2.5%. With the further addition of lecithin, the serum requirement for maximum granulocyte/macrophage colony growth was reduced from 10 to 2.5%. Again it was interesting that whereas erythroid colony formation was dependent on serum even in the presence of these substances, some granulocyte/macrophage colony formation was possible in the absence of serum.

Table 2 Effect of BSA, Tf, Se and egg lecithin (lec) on granulocyte–macrophage colony counts

Addition		Colonies per 10⁵ nucleated cells
15% FCS,		146 ± 9*
None		0
Se		0
BSA		0
Tf		0
Lec		0
Se,	BSA, Tf, lec	57 ± 4
— ,	BSA, Tf, lec	46 ± 7
Se,	— , Tf, lec	0
Se,	BSA, — , led	0
Se,	BSA, Tf, —	24 ± 3

BSA was used in a concentration of 1.2%.
Concentrations of Tf and Se were as for Fig 1.
*Standard error.

It is possible that the activities demonstrated for the proteins BSA and transferrin were attributable to trace impurities. Efforts were made to diminish this possibility by further purification of supplied materials (legend Fig. 1). Transferrin activity was not replaced by the equivalent amount of Fe^{3+} alone, nor was its activity reduced after gel permeation chromatography. Acrylamide slab gel electrophoresis[8] of purified BSA and transferrin in denaturing conditions (0.1% SDS, 0.1% β-mercaptoethanol, 2 µg of protein per slot) revealed only a single Coomassie brilliant blue-stainable band in both instances.

We anticipate that the active substances reported here will prove to have similar nonspecific, serum-replacing activity in cultures of various other cells and tissues. Selenium, present in serum at a concentration of

1.4×10^{-7} M (ref. 9), has long been recognised as an essential trace element in whole animal studies[10], and McKeehan et al. demonstrated it to be a requirement for growth of human fibroblasts in culture[11]. The selenium-containing enzyme glutathione peroxidase is known to be present in leukocytes[12] as well as erythrocytes[13]. The requirement for transferrin by maturing red cells is not surprising, since this iron-transporting protein, present in serum at 4×10^{-5} M (ref. 14), is known to be the major source of iron for these cells. It was unexpected, however, that transferrin should also prove necessary for growth of granulocyte/macrophage precursors. Taken together with the observations that transferrin has a role in the proliferative response of human lymphocytes to PHA[15], and in the growth of fibroblasts and pituitary cells in culture[16], the finding suggests a much wider requirement for this protein than has been recognised. Among numerous conceivable functions of albumin, we consider the likeliest to be that of a buffer for components of the medium present in inhibitory amounts. This suggestion predicts that the albumin requirement would diminish if the numerous defined components of the culture medium were adjusted to optimum levels.

The straightforward interpretation of our observations is that a major role of serum in culture is to provide selenium, transferrin, albumin and lipids to the medium. Our results do not, however, exclude the possibility that these entities simply substitute for other unrelated substances provided by serum. Several agents that were expected to show activity did not enhance colony formation at low serum concentrations, provided that selenite, transferrin, albumin and lecithin were present and that erythropoietin and colony-stimulating factor were maintained at non-limiting levels. Among those tested were dexamethasone (2.5×10^{-10} M), insulin (2.5×10^{-8} M), testosterone (10^{-8} M), etiocholanolone and fluoxymesterone (3×10^{-7} M), prostaglandins E_1, E_2 and $F_2\alpha$ (5×10^{-9} M), $ZnCl_2$ (10^{-7} M), and cholesterol (3×10^{-7} M).

Our observations have important technical implications. First, inclusion in culture medium of the active substances reported here greatly reduces the quantities of serum required, and in our experience has transformed several serum lots from inadequate to fully active. More important, we expect their routine inclusion in medium to simplify the search for more specific biologically active substances in conditioned media and serum by reducing the number of variables in the system.

L. J. Guilbert
N. N. Iscove

Friedrich Miescher-Institut,
PO Box 273,
CH-4002 Basel, Switzerland

1 Stephenson, J. R., Axelrad, A. A., McLeod, D. L., and Shreeve, M. M., *Proc. natn. Acad. Sci. U.S.A.*, **68**, 1542–1546 (1971).
2 Iscove, N. N., Sieber, F., and Winterhalter, K. H., *J. cell. Physiol.*, **83**, 309–320 (1974).
3 Metcalf, D., *Expl. Hemat.*, **1**, 185–201 (1973).
4 Iscove, N. N., and Sieber, F., *Expl. Hemat.*, **3**, 32–43 (1975).
5 Bradley, T. R., and Sumner, M. A., *Aust. J. exp. Biol. med. Sci.*, **46**, 607–618 (1968).
6 Worton, R. G., McCulloch, E. A., and Till, J. E., *J. cell. Physiol.*, **74**, 171–182 (1969).
7 Armelin, H. A., Nishikawa, K., and Sato, G. H., in *Control of Proliferation in Animal Cells* (edit. by Clarkson, B., and Baserga, R.), 97–104 (Cold Spring Harbor Laboratory, New York, 1974).
8 Studier, F. W., *Science*, **176**, 367–376 (1972).
9 *Biology Data Book*, second ed., **3**, (edit. by Altman, P. L., and Dittmer, D. S.), 1751 (Federation of American Societies of Experimental Biology, Bethesda, 1974).
10 West, E. S., and Todd, W. R., *Textbook of Biochemistry*, third ed., 1262–1263 (MacMillan, New York, 1961).
11 McKeehan, W. J., Hamilton, W. G., and Ham, R. G., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 2023–2027 (1976).
12 Reed, P. W., *J. biol. Chem.*, **244**, 2459–2464 (1969).
13 Mills, G. C., *J. biol. Chem.*, **229**, 189–197 (1957).
14 Bezkorovainy, A., and Zschocke, R. H., *Drug Res.*, **24**, 476–485 (1975).
15 Tormey, D. C., Imrie, R. C., and Mueller, G. C., *Expl. Cell Res.*, **74**, 163–169 (1972).
16 Hayashi, I., and Sato, G. H., *Nature*, **259**, 132–134 (1976).

Cell contact induces an increase in pinocytotic rate in cultured epithelial cells

ENDOCYTOSIS is the process by which cells internalise part of the plasma membrane in association with extracellular components: they engulf large particles by phagocytosis and internalise components in the fluid phase by pinocytosis. Discrimination of phagocytosis from pinocytosis has been based on the size of the particle and the amount of energy required to internalise it, phagocytosis being inhibited more readily by metabolic poisons than pinocytosis[1-3]. This is consistent with the view that more energy is required to internalise a large surface area than is needed for a small surface area. Vasiliev et al.[4] described "contact inhibition of phagocytosis": when cultured epithelial cells, but not fibroblasts, became confluent, the rate of phago-cytosis decreased considerably. I report here that confluent cultured epithelial cells are not inhibited in their ability to undergo pinocytosis, but exhibit a greater rate of pinocytosis than do non-confluent cells. This difference is not due to diffusable factors, but is the result of cell contact.

Permanent lines of cultured epithelial cells of human (HeLa) or monkey origin (Mk$_2$) were grown as monolayers on plastic tissue culture dishes and maintained in Eagle's MEM containing 10% calf serum, penicillin (5 U ml^{-1}) and streptomycin (2 U ml^{-1}). Cells were passaged routinely every 4-5 d. Pinocytosis was measured using horse-radish peroxidase (HP) according to the method of Steinman et al.[5,6], or ^{125}I-albumin by Rhyser's procedure[7]. HeLa cells took up HP linearly with respect to both concentration (not shown) and time (Fig. 1). The level of endogenous peroxidase activity in HeLa cells was essentially unmeasurable (<0.1 ng enzyme per mg cell protein), and washing removed all non-internalised enzyme. These results support the conclusions of Steinman et al.[5,6] on the value of HP as a pinocytotic marker. I also observed that when the uptake of enzyme was calculated on either a per protein (Fig. 1a) or a per cell basis (not shown), cultures of HeLa cells plated at higher densities exhibited greater rates of uptake than those of lower densities. In spite of slight differences in the

absolute amount of HP taken up in different experiments (10-20%), there was a consistent difference in the rate of uptake between cultures seeded at different densities. This is demonstrated more strongly in Fig. 2, in which a wider range of cell densities was used. Similar differences in the rate of uptake were observed whether marker was presented to cells in fresh medium or medium conditioned by exposure to confluent cells for 3 d.

Mk$_2$ cells are derived from a pool of normal rhesus monkey kidneys, and unlike HeLa cells, they have several characteristics typical of normal cells, such as density-dependent inhibition of growth and macromolecular synthesis, and an inability to grow in either suspension or soft agar[8]. I also observed that as Mk$_2$ cultures became more dense, the rate of pinocytosis increased, and stabilised only as they reached their saturation density (data not shown). Similar results were obtained if pinocytosis was assessed on a per cell instead of a per protein basis. Thus, a density-dependent increase in the rate of pino-cytosis is not unique to HeLa cells.
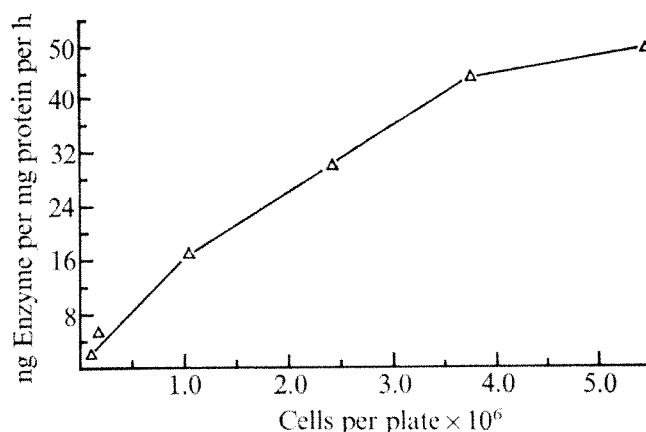
Fig. 2 Uptake of HP by different densities of HeLa cells. Different numbers of HeLa cells were seeded on plates 72 h before the experiment. Monolayers were incubated in medium containing HP (1 mg ml^{-1}) for 3 h and then processed as described for Fig. 1. Duplicate samples were taken and cells were counted.

The difference in rate of uptake was independent of HP. When iodinated albumin was used as a pinocytotic marker, cultures of HeLa or Mk$_2$ cells at a higher density exhibited a greater rate of uptake than those at a lower density.

To establish that the difference in accumulation of pinocytotic markers was the result of a difference in uptake, I measured the rate of intracellular inactivation of HP. Cultures of HeLa cells plated at different densities and exhibiting a twofold difference in uptake inactivated the enzyme at the same rate (Fig. 3). In both cases, the enzyme was inactivated with single exponential kinetics (correlation coefficient >95%), yielding a half life of 6.5 h. HeLa cells plated at different densities degrade ^{125}I-albumin at similar rates, yielding a half life of 16.4 h. Thus, differences in the rate of accumulation reflect differences in the rate of uptake rather than degradation.

The progressive increase in the rate of pinocytosis as cell density increased (Fig. 2) was independent of cell number but not cell density as the following experiment showed. Using small glass cylinders of different diameter, I plated a constant number of cells per dish. This procedure yields different cell densities, thus modifying the degree of cell-cell contact. I found that an increase in cell density gave rise to an increase in pino-cytosis regardless of whether the pinocytotic marker was presented to cells in fresh medium or medium conditioned by exposure to confluent monolayers. Thus, an increase in the degree of cell contact induced an increase in the rate of pino-cytosis.

As the difference between rates of uptake by cells plated at different densities was the same in fresh and conditioned

Fig. 1 Cultures of different numbers of HeLa cells were seeded 48 h before the experiment. Cells were then incubated in complete medium containing HP (1 mg ml^{-1}) for the times specified. After incubation the monolayers were washed eight times with cold phosphate-buffered saline (PBS), incubated in PBS at 4 °C for 20 min and washed four more times. The monolayers were then solubilised with 0.1% Triton X-100 and HP was assayed as before[5]. Duplicate cultures were trypsinised and cells were counted. Protein concentrations were determined by the procedure of Lowry et al.[10]. Each point is the average of triplicate samples. The standard error of the mean is 5-8% for both cell number and enzyme assay. $\triangle$, 1.25 × 10^6 cells per plate; $\bigcirc$, 0.75 × 10^6 cells per plate.

Fig. 3 Rate of degradation of internalised HP. Monolayers of cells were incubated with HP (1 mg ml^{-1}) for 20 h and then washed six times with cold MEM and twice with warm MEM + 10% calf serum. At the times specified, monolayers were collected and enzyme was assayed as described for Fig. 1. The symbols represent experimental data, the line is a computer-determined best fit to the data. At the end of the experiment the cultures had densities of 1.41 ($\triangle$) and 3.81 × 10^6 ($\bigcirc$) cells per plate.

medium, I suggest that diffusable factors are not responsible for the difference in the rate of pinocytosis. I found that if cellular levels of cyclic AMP were increased by addition of dibutyryl cyclic AMP (1 mM) or caffeine (10 mM), the rate of pinocytosis in sparse cultures was not affected.

The change in rate of pinocytosis in HeLa and Mk$_2$ cells occurs before any overt change in growth rate. For example, Mk$_2$ cells divide at a constant rate with a generation time of 42 h until a density of 5.5 × 10^5 cells per cm^2 is reached[8]. HeLa cells divide at a constant rate until a cell density of 1.2 × 10^6 cells per cm^2 is reached, after which the rate of division decreases by 80–90% (my unpublished results). But before any change in growth rate, the rate of pinocytosis may change by a factor of four to six. This suggests that changes in the rate of pinocytosis are independent of inhibition of cell division.

The cell types used in this study were of epithelial origin. It has been demonstrated, however, that L cells[6] and sarcoma 180 cells (H.J.P. Rhyser, personal communication) of fibroblast origin exhibit a density-induced increase in pinocytosis. Apart from the Mk$_2$ cell all these cells are derived from established tumours. Although Mk$_2$ has several characteristics typical of normal cells, it has been maintained in culture for 16 yr; perhaps the density-dependent increase in pinocytosis results from adaptation to tissue culture. I am investigating whether primary cultures of epithelial cells exhibit a density-dependent increase in pinocytosis.

Using either polystyrene beads (1.10 µm in diameter) or Oil red 0–albumin emulsions, I have not been able to quantify the rate of phagocytosis in either non-confluent or confluent monolayers incubated for up to 4 h. Thus I cannot comment directly on "contact inhibition of phagocytosis". It may be that the difference in phagocytosis is related to a change in the adhesive characteristics of the membranes of confluent cells, as suggested previously[4,9]. Whether or not "contact inhibition of phagocytosis" can be confirmed, my results show that it is not due to the inability of the cell to internalise plasma membrane. Contrary to expectation, the occurrence of cell–cell contact seems to enhance at least one membrane function.

I thank Ms M. Nielson for technical assistance and Ms J. Morneault for typing the manuscript. This work was supported

Department of Physiology,
University of Connecticut Health Center,
Farmington, Connecticut 06032

1 Karnovsky, M. L., *Physiol. Rev.*, **42**, 143–168 (1962).
2 Cohn, Z. A., *J. exp. Med.*, **124**, 555–571 (1966).
3 Petitpierre-Gabathuler, M. P., and Rhyser, H. J-P., *J. Cell Sci.*, **19**, 141–156 (1975).
4 Vasiliev, S. M., Gelfand, I. M., Domina, L. U., Zacharova, D. S., and Jubimov, L., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 719–722 (1975).
5 Steinman, R. M., and Cohn, Z. A., *J. Cell Biol.*, **55**, 186–204 (1972).
6 Steinman, R. M., Silver, J. M., and Cohn, Z. A., *J. Cell Biol.*, **62**, 949–969 (1974).
7 Rhyser, H. J. P., *Science*, **159**, 390–396 (1968).
8 Kaplan, J., and Moskowitz, M., *Biochim. biophys. Acta.*, **389**, 290–305 (1974).
9 DiPasquale, X. X., and Bell, P. B., Jr, *J. Cell Biol.*, **62**, 198–214 (1974).
10 Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J., *J. biol. Chem.*, **193**, 652–275 (1951).

Presence of membrane particles in freeze-etched bovine olfactory cilia

WHEN an odorant molecule stimulates an olfactory receptor cell the electrical changes which are set up must result from some kind of interaction between the stimulant and the cell surface. Various theories of olfaction have favoured either lipids[1] or proteins[2], as the principal site of odorant reception. Experiments distinguishing unequivocally between these alternatives have yet to be performed[3], but there is growing evidence that the receptor is proteinaceous[4,5].

Two other protein-containing receptors that have been extensively explored are rhodopsin[6] and the acetycholine receptors of skeletal muscle[7] and electroplax origin[8]. All of these have been visualised in freeze-fractured preparations by electron microscopy as being related to particles 8–12 nm

Fig. 1 Freeze-etch replica of the surface of the bovine olfactory epithelium showing two receptor endings (re) and the apex of a supporting cell (sc). Note the broken basal portions of cilia at the surface of the sensory ending on the right, and two vertically cleaved cilia on the left (× 25,000).

Fig. 2 Detail of Fig. 1 showing membrane features of two sensory cilia. In the cilium on the right, numerous membrane particles and part of a ciliary necklace (arrow) are visible on the exposed PF face. In contrast the EF face shown in the cilium to the left, bears few or no membrane particles ($\times$90,000).

across, embedded in the cell membranes of the sensory surfaces. If the olfactory receptor sites are indeed of a similar nature, one also might expect to find membrane particles in the sensory cilia of olfactory receptor cells, but until now attempts to observe them have been unsuccessful[9-11]. In this letter we report the discovery of high densities of membrane particles in the sensory cilia of bovine olfactory cells.

Heads of male calves (1 month old) were longitudinally bisected and olfactory tissue on the cribiform plate and ethmoturbinates fixed *in situ* with ice-cold Karnovsky's fixative[12] buffered at pH 7.0, for 1 h, then excised and placed in fresh fixative for 2 h. Tissue from the neighbouring ciliated non-sensory respiratory mucosa was also treated in the same way for control purposes. After washing in buffer, small pieces of tissue were soaked in 20% cacodylate-buffered glycerol (pH 7.0) containing 0.01% $CaCl_2$ for 24 h and were then fractured, etched and replicated in a Balzers GA-6 freeze etching device according to standard techniques[13]. Replicas were observed in an Hitachi 125E electron microscope at 125 kV. The nomenclature recently proposed by Branton *et al.*[14] will be used in this description. In the figures, shadows appear white.

In addition to studies of freeze-etched material, pieces of tissue were also examined in sections by routine transmission and scanning electron microscopy for comparison with the freeze-etched material. In this way, the olfactory and non-olfactory tissue could be unequivocally distinguished in both types of preparation.

The general structure of the sensory tissue of the ox is similar to that of other mammals which have been studied with the electron microscope[15]. The surface of the olfactory epithelium is formed partly by the terminal knobs of sensory dendrites bearing an average of 17 cilia which possess short proximal portions averaging 1.7 μm long by 0.2 μm wide, tapering to long trailing distal regions 0.08 μm in diameter and at least 30 μm long; these distal portions are expanded

occasionally to form vesicles averaging 0.6 μm across. The supporting cells surround and separate the receptor cells, and bear at their apices abundant microvilli which extend almost to the surface of the mucus layer covering the epithelium, the thickness of this layer being about 5 μm. The neighbouring respiratory epithelium consists mainly of non-sensory columnar epithelial cells bearing motile cilia and a few microvilli, with occasional goblet cells interspersed.

Freeze fractured and etched olfactory epithelia showed a pattern of recognisable sensory endings and cilia, and the surfaces of supporting cells. Membrane particles, 8–12 nm across, were present in abundance on the olfactory cilia, at the PF-face (A face) of the membrane (Figs 1–3), and over the whole of the exposed sensory surface of the dendrite. Particles were more numerous on the narrow distal ends of the cilia (6,000 μm^{-2}) than on the more proximal ends (4,300 μm^{-2}) or on the dendrite ending itself (3,400 μm^{-2}); the mean density for the surface of the sensory ending was 5,200 μm^{-2}.
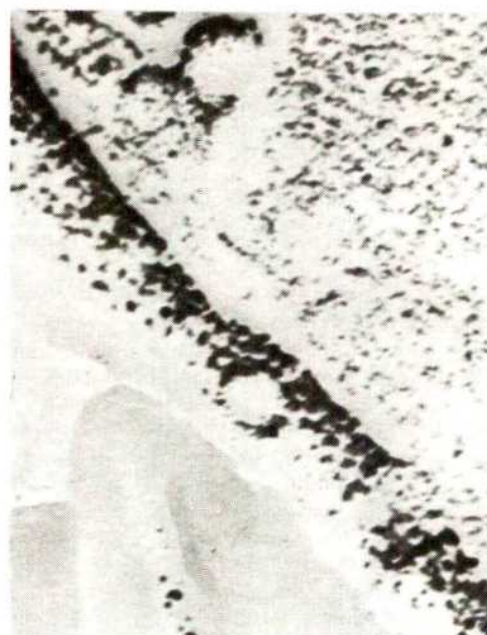
Fig. 3 The tapering distal portion of an olfactory cilium covered in closely packed membrane particles. The ring-like structure in the centre is a surface irregularity ($\times$190,000).

From measurements made on transmission and scanning electron micrographs it was calculated that each sensory ending had a surface area of at least 160 μm^2, bearing approximately 10^6 particles. The number of endings per cm^2 of epithelial surface was 5$\times$10^6. Allowing for the volume occupied by receptor and supporting cell processes, the concentration of particles within the confines of the 5 μm layer of mucus at the epithelial surface was estimated to be 10^{19} l^{-1}, which represents a molarity of 2$\times$10^{-5} if each particle is assumed to be a single molecule.

Close to the proximal end of each olfactory cilium, six rows of particles, corresponding to the "necklace" formation in other cilia[16], could be distinguished. In contrast the cilia of the neighbouring non-sensory respiratory epithelium bore few membrane particles (average 300 m^{-2}) although the proximal 'necklace' was present, as described by other authors[17] (Fig. 4). Membrane particles were, however, seen on the short microvilli interspersed among these cilia. Few particles were found at the EF-face (B face) of the membranes of either the olfactory or non-olfactory tissues.

The particles found in the sarcolemma at neuromuscular junctions and in the rod outer segment are similar in packing density to those described here[6,7]. Membrane particles have been shown in several cases to be proteinaceous[18] and

10 Stockinger, L., Kerjaschki, D., and Hörandner, H., *Verh. anat. Ges., Jena,* **68,** 487–490 (1974).
11 Altner, H., and Altner, I., in *Transduction Mechanisms in Chemoreception* (edit. by Poynder, T. M.), 59–70 (International Retrieval Ltd, London, 1974).
12 Karnovsky, M. J., *J. Cell Biol.,* **27,** 137A–138A (1965).
13 Fisher, K., and Branton, D., *Meth. Enzymol.,* **32,** 35–44 (1975).
14 Branton, D., and others, *Science,* **190,** 54–56 (1975).
15 Graziadei, P. P. C., in *The Ultrastructure of Sensory Organs* (edit. by Friedman, I. F.), 267–305 (North Holland, Amsterdam, 1973).
16 Gilula, N. B., and Satir, P., *J. Cell Biol.,* **53,** 494–509 (1972).
17 Sleytr, U., Kerjaschki, D., and Stockinger, L., *Mikroskopie,* **28,** 257–268 (1972).
18 Branton, D., *Phil. Trans. R. Soc.,* **B261,** 133–138 (1971).
19 Gesteland, R. C., Lettvin, J. Y., and Pitts, W. H., *J. Physiol., Lond.,* **181,** 525–559 (1965).
20 Kerjaschki, D. and Hörandner, H., *J. Ultrastruct. Res.,* **54,** 420–444 (1976).

Fig. 4 The basal region of a non-sensory cilium from the respiratory epithelium. Note the scarcity of membrane particles distal to the necklace ($\times 184,000$).

it can be predicted that the olfactory ciliary membranes are appreciably richer in protein than, for example, those of the motile respiratory cilia. Although the particles could represent ion conductance channels, or some purely structural feature, it seems reasonable to suppose they might also be the receptor sites responsible for binding odorants. Their presence in such high densities may be correlated with the sensitivity of the sensory cells to low concentrations of odorants, and also with the hypothesis that many classes of receptor site, responding to different odorants, may be present on individual sensory cells[19].

Note added in proof: Similar particles have recently been found in mouse olfactory cilia[20].

B. P. M. Menco*
G. H. Dodd

*Department of Molecular Sciences,
University of Warwick,
Coventry CV4 7AL, UK*

M. Davey

*Department of Botany,
The University,
Nottingham NG9 3DR, UK*

L. H. Bannister

*Departments of Biology and Anatomy,
Guy's Hospital Medical School,
London SE1 9RT, UK*

1 Davies, J. T., *Symp. Soc. exp. Biol.,* **16,** 170–179 (1962).
2 Amoore, J. E., in *Taste and Smell in Vertebrates, Ciba Symposium* (edit. by Wolstenholme, G. E. W., and Knight, J.), 293–312 (Churchill, London, 1970).
3 Dodd, G. H. in *Transduction Mechanisms in Chemoreception* (edit. by Poynder, T. M.), 103–113 (International Retrieval Limited, London, 1974).
4 Getchell, M. L., and Gesteland, R. C., *Proc. natn. Acad. Sci. U.S.A.,* **69,** 1494–1498 (1972).
5 Gennings, J. N., Gower, D. B., and Bannister, L. H., *J. Endocr.,* **67,** 47P–48P (1975).
6 Jan, L. Y., and Revel, J.-P., *J. Cell Biol.,* **62,** 257–273 (1974).
7 Hartzell, H. C., and Fambrough, D. M., *J. gen. Physiol.,* **60,** 248–262 (1972).
8 Orci, L., Perrelet, A., and Dumant, Y., *Proc. natn. Acad. Sci. U.S.A.,* **71,** 307–310 (1974).
9 Kerjaschki, D., Sleytr, U., and Stockinger, L., *Naturwissenschaften,* **59,** 314–315 (1972).

Inhibition of PEP-carboxykinase in rat liver by polychlorinated biphenyl

POLYCHLORINATED biphenyls (PCBs) are widely used industrial chemicals which have found their way into the environment, where they have been detected in rainwater, in many species of birds and fish and in human tissues. PCBs accumulate in the body's fat tissues, where they apparently degrade very slowly under natural conditions[1], but recently it has been found[2] that PCBs may act as carcinogens. The detailed actions of PCBs inside the body, such as their effects on carbohydrate, lipid or protein metabolism, are relatively unknown although they have been shown to alter the liver cell in mouse and monkey[3], to affect the immune system[4] and to alter the activity of microsomal and other enzyme systems[5–8]. We have studied the *in vivo* effect of PCBs on the activity of two regulatory
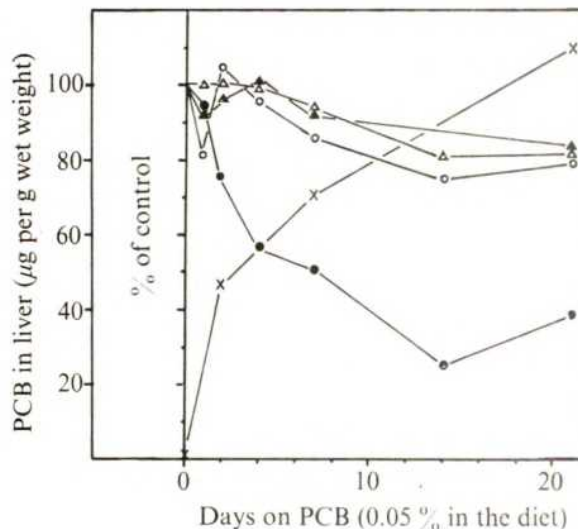
Fig. 1 Effects of PCB (Clophen T64) on PEPCK and FDPase activities in rat liver and on the level of blood glucose. Female rats (Sprague–Dawley, 150–170 g) were used in the experiments. PCB diet was prepared by thoroughly mixing (starmix) 500 g of normal pellets (Eggersmann, Rinteln) with 10 ml of a solution of PCB (Clophen T64) in ethanol giving a final concentration of 0.05% of PCB in the diet. The powder was pasted with water, pellets were formed and dried for 2 d at 37 °C in an oven. The diet for control animals was treated in exactly the same way but without PCB. After the indicated time intervals, livers of 4 rats from each group were removed under ether anaesthesia, homogenised individually in a Potter homogeniser at 0 °C for 1 min using 2 g of liver tissue per 10 ml of 0.25 M sucrose solution. The homogenate was centrifuged for 30 min at 105,000g. PEPCK activity[11] and FDPase activity[12] were determined in the supernatant. Blood (0.10 ml) was taken from the aorta descens during removal of the liver, and blood glucose was determined with glucose oxidase[13]. Enzyme activities and content of blood glucose from control animals were set at 100%. $\times$, PCB in liver; $\blacktriangle$, food intake; $\triangle$, blood glucose; $\bigcirc$, FDPase; $\bullet$, PEPCK.

enzymes of gluconeogenesis, phosphopenol-pyruvate carboxykinase (PEPCK) and fructose-1, 6-diphosphatase (FDPase), respectively, in rat liver.

As can be seen from Fig. 1, a diet containing 0.05% PCB (Clophen T64, 60% Cl) diminished the activity of PEPCK by about 50% within a feeding period of 7 d. FDPase activity and the content of blood glucose were also influenced by PCB, but to a much lower extent. Furthermore, the food intake of the animals was not markedly altered by PCB when given in this concentration. During the experimental period the content of PCB in the liver tissue increased from 0.145 (control) to 110 µg per g wet liver. Thus, an *in vivo* content of 70.5 p.p.m. of PCB in the liver lowered the activity of PEPCK to one half of its original activity.

Table 1 Effects of various chlorinated compounds on the activities of PEPCK and FDPase and on the content of blood glucose in rat liver

Dietary addition (0.05%, 7 d)	PEPCK	FDPase	Blood glucose
		(% of control)	
None	100	100	100
Biphenyl	91	101	79
2-Chlorbiphenyl (19% Cl)	119	102	104
Clophen A30 (42% Cl)	66	110	97
Clophen A50 (54% Cl)	53	75	78
Clophen A60 (60% Cl)	51	84	99
Hexachlorobenzene	106	96	100

Rats were fed a normal diet (control), a diet with biphenyl (0.05%) or with various chlorinated compounds (0.05%) for 7 d. Preparation of the diet and experimental details are given in the legend to Fig. 1.

Figure 2 shows the effect of increasing concentrations of PCB in the food after a feeding period of 4 d. A dose of 0.01–0.05% PCB reduced the activity of PEPCK, but did not alter the activity of FDPase, the level of blood glucose and the food intake. Higher concentrations of PCB, however, were obviously toxic because the food intake was reduced and the animals partially fasted. Consequently, PEPCK activity increased in these conditions, since the activity of this enzyme is very sensitive to the animals' nutritional state (see Table 2 and ref. 9). The level of blood glucose and FDPase activity were also lowered by higher concentrations of PCB.

Other chlorinated compounds which are chemically related to PCBs or are concomitants (hexachlorobenzene) were also tested for their *in vivo* ability to influence the activity of PEPCK, FDPase and the level of blood glucose (Table 1). From these compounds PCBs with high chlorine content were most effective on PEPCK, whereas the effect on FDPase activity and on the level of blood glucose varied.

Gluconeogenic conditions such as fasting or feeding a protein diet (poor in carbohydrate) increases PEPCK activity[9]. As is shown in Table 2, after fasting or protein diet PEPCK activity was increased 4.5-fold and threefold, respectively. When PCB was added to the normal diet, the induction rate after
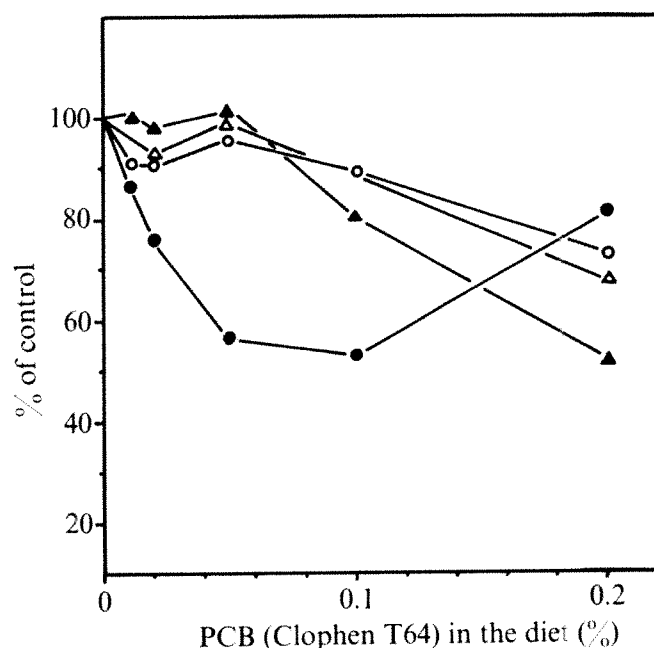
Fig. 2 Effect of increasing PCB concentration on PEPCK and FDPase activities in rat liver. Rats were fed a diet with PCB (Clophen T64) for 4 d. Diet was prepared as described in the legend to Fig. 1, but PCB concentration was varied. Further details see legend to Fig. 1. Symbols as in Fig. 1.

fasting was only twofold and in rats fed a protein diet containing PCB the increase of PEPCK activity was only 1.8-fold. Thus, PCB severely disturbs the ability of the rats to adapt PEPCK activity to gluconeogenic conditions.

Although PCBs are chemically almost inert substances, these compounds nonetheless can alter intermediary metabolism, that is gluconeogenesis. Most determinations described here were carried out between 4 and 7 d on PCB diet (Fig. 2 and Table 1, respectively). This corresponds to an *in vivo* PCB concentration in the liver of about 55–70 µg PCB per g wet liver tissue (see Fig. 1), being in the range of the concentration of physiological substrates. One can suggest, therefore, that PCBs do not act on PEPCK itself but somehow affect its regulation. Since PEPCK activity is closely connected with the level of cyclic AMP and hormones[10], these would be good candidates for further investigations on the mechanism of action of PCBs on intermediary metabolism.

We thank W. Sümmermann and N. Schwaak for the determinations of PCB contents in liver tissue.

B. Messner
J. Berndt
J. Still

*Institut für Biochemie der Gesellschaft
für Strahlen- und Umweltforschung,
Zentralabteilung, 8 München 2,
Landwehrstrasse 61, FRG*

Table 2 PEPCK induction after PCB administration

Treatment of rats	PEPCK	FDPase	Glucose
	(specific activity nmol mg^{-1} min^{-1})		(mg per 100 ml)
Fed normally	57.8 ± 4.4	34.5 ± 1.5	165 ± 12
Fasted for 16 h (overnight)	260.2 ± 14.8	32.4 ± 1.4	86 ± 3
Clophen T64 (0.05%, 4 d) + fasted for 16 h	121.0 ± 14.3	24.2 ± 2.6	85 ± 12
Protein diet (4 d)	179.6 ± 4.5	27.8 ± 1.2	166 ± 28
Protein diet* with Clophen T64 (0.05%, 4 d)	106.2 ± 5.2	26.7 ± 1.0	121 ± 10

*Protein diet was from Eggersmann, Rinteln. For preparation of PCB-diet and experimental details see legend to Fig. 1.

1 Hammond, A. L., *Science*, **175**, 155–156 (1972).
2 Kimbrough, R., *New Scient.*, **68**, 331–336 (1975).
3 Nishizumi, M., *Archs Environ. Hlth*, **21**, 620–632 (1970).
4 Vos, J. G., and De Roij, Th., *Tox. appl. Pharmac.*, **21**, 549–555 (1972).
5 Litterst, C. L., Farber, T. M., Baker, A. M. and van Loon, E. J., *Tox. appl. Pharmac.*, **23**, 112–122 (1972).
6 Villeneuve, D. C., Grant, D. L., Phillips, W. E. J., Clark, M. L., and Clegg, D. J., *Bull. environ. Contam. Tox.*, **6**, 120–128 (1971).
7 Alvares, A. P., Bichers, D. R., and Kappas, A., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 1321–1325 (1973).
8 Davis, P. W., Friedhoff, J. M., and Wedemeyer, G. A., *Bull. environ. Contam. Tox.*, **8**, 69–72 (1972).
9 Lardy, H. A., Foster, D. O., Shrago, E., and Ray, P. D., *Adv. Enzyme Reg.*, **2**, 39–47 (1964).
10 Krone, W., Huttner, W. B., Seitz, H. J., and Tarnowski, W., *FEBS Lett.*, **52**, 85–89 (1975).
11 Seubert, W., and Huth, W., *Biochem. Z.*, **343**, 176–191 (1965).
12 Racker, E., *Meth. Enzymol.*, **5**, 272–273 (1962).
13 Bergmeyer, H. U., and Bernt, E., in *Methoden der enzymat. Analyse*, 3rd ed. (edit. by Bergmeyer, H. U.), 1250–1258 (Verlag Chemie, 1974).

Two-gene control of T-helper cell induction

THE major histocompatibility complex (MHC) of mouse, man and other species has attracted much attention recently, and that of the mouse is perhaps the most precisely studied mammalian genetic region (reviewed in refs 1–3). Within the space of 0.5 centimorgans are genes coding for transplantation antigens present on essentially all cells (H-2K and H-2D); genes influencing immune responses ("immune response or *Ir* genes"); genes coding for surface alloantigens of restricted tissue distribution (immune response gene region associated antigens or Ia antigens); genes influencing cell interactions, the production of factors and stimulation in the mixed lymphocyte reaction (all in the I region and its subregions I-A, I-B, I-C and so on), as well as genes controlling serum proteins related to complement (S region), and red-cell antigens. Although the products of some of these genetic regions are clear (for example, of the H-2D or H-2K loci) this is not the case for the I region which influences many immune functions, such as delayed hypersensitivity[4] and antibody production[5]. One group of products of the I region genes constitutes the Ia antigens which have been defined by alloantisera[6,7]. Other products of these genes are still elusive, and some authors have even suggested that I region gene products are antigen-specific receptors on T cells[8,9].

Recently, it was found that the I region influenced the induction of helper cells, *in vitro*[10]. To induce T-helper cells in our *in vitro* system[11], macrophages are required[12,13], and with soluble antigens the interaction between T cells and macrophages required genetic similarity in the I-A subregion of the H-2 complex[10]. In the mouse, unlike the guinea pig[14] interaction did not depend on direct contact between T cells and macrophages[12], and was mediated by at least two factors[10]. (1) A genetically related macrophage factor (GRF) is obtained from antigen-treated macrophages and generates T-helper cells in the absence of macrophages or additional antigen, but only if the macrophage releasing GRF and the T cells are similar at the I-A subregion of the H-2 complex. (2) A nonspecific macrophage factor (NMF), obtained from macrophages incubated without antigen, is not genetically restricted and induces T-helper cells in the absence of macrophages but only with particulate antigens[10]. We have characterised the immunochemical nature of GRF and found that it consists of I
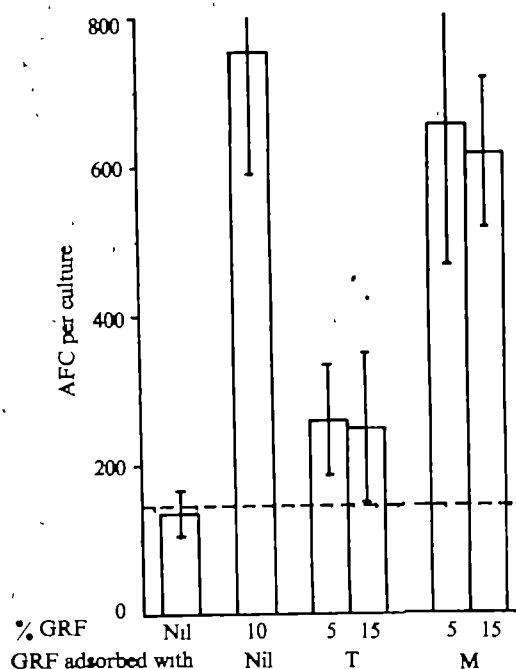
Fig. 1 Evidence for a receptor for GRF on lymphoid cells but not on macrophages. CBA GRF$_{KLH}$ (1 ml) was incubated with either 1×10^8 nylon wool-purified T cells from CBA spleen (T) or with 6×10^7 macrophages (M) (purified by irradiation with 2,000R and adherence to plastic) for 1 h at room temperature. After removal of the cells by centrifugation the adsorbed GRF were added to 15×10^6 nylon wool-purified CBA T cells in a concentration of either 5 or 10% and incubated as described in the legend to Table 1. After 4 d, 2×10^6 living cells from each culture were added to 15×10^6 CBA spleen cells and incubated with TNP-KLH (0.1 µg ml^{-1}). Four days later the anti-DNP-response was measured. The dotted line represents the background (15×10^6 spleen cells and TNP-KLH, but no helper cells added). Five experiments of this type have been performed with analogous results. The vertical bars represent the standard error of the arithmetical mean of three independent cultures. To exclude the presence of receptors on B cells, anti-θ-treated and complement-treated spleen cells, cleared of dead cells by low ionic strength buffer[20] were used (viability >95%); 10^6 B cells per ml did not remove GRF (for example, in a typical experiment CBAT+GRF: 597 AFC per culture, GRF adsorbed B cells: 480, GRF adsorbed T cells: 47).

Table 1 GRF contains I-A subregion-coded products

CBA GRF$_{KLH}$	Immunoadsorbents		S/N%	Anti-DNP-response (AFC±s.e.) first eluate/acid eluate	
−	—		—	18±16	
+	—		10	332±37	
+	A.TL anti-A.AL	(aK^k)	5	250±27	30±17
+			15	220±68	0
+	(A.BY × B10.HTT)F$_1$ anti-A.TL	(aI-A^k)	5	0	363±46
+			15	0	400±56
+	B10.A(4R) anti-B10.A(2R)	(aI-B^k)	5	199±26	ND
+			15	252±36	ND
+	(B10.A(4R) × 129)F$_1$ anti-B10.A(2R)	(aI-C$^{k\,\cdot\,d}$)	5	423±112	0
+			15	373±23	27±12

GRF obtained from CBA macrophages by incubating with keyhole limpet haemocyanin (KLH, 10 µg per 5×10^6 cells) for 4 d (CBA GRF $_{KLH}$) were passed over columns containing Sepharose 2B (1 ml of GRF per ml of Sepharose) conjugated with the immunoglobulin fractions of antisera directed against the K^k, I-A^k, I-B^k or I-C^k or ^{d} subregion of the H-2 complex (first eluate). After washing, the immunoadsorbents were treated with 1 ml of Sørensen's glycine buffer, pH 2,4; the eluates were neutralised immediately and dialysed against phosphate-buffered saline, pH 7.2 (acid eluate). 'First and acid eluates' were tested in different concentration for their helper-cell-inducing capacity by incubating with 15×10^6 nylon wool-purified CBA T cells for 4 d in Marbrook type flasks as described in detail elsewhere[11,12,13]. Living cells (2×10^6) from each culture (which include helper cells) were then added to 15×10^6 CBA spleen cells and trinitrophenylated KLH(TNP-KLH, 0.1 µg ml^{-1}), and 4 d later the anti-dinitrophenyl (DNP)-response was measured by the plaque forming ability of the B cells against DNP or TNP (which cross react) coated sheep red blood cells (SRBC). The results are given as arithmetic means of the DNP-AFC (antibody forming cells)±s.e. for three independent cultures. Specific AFC are enumerated by subtracting the number of plaques obtained with SRBC from that obtained with DNP-SRBC. The controls include 15×10^6 spleen cells incubated with either TNP-KLH (28±17), or DNP-polyacrylamide beads 3% (1372±175) or no antigen (12±11). (B10.A(4R) × 129) F, anti-B10.A(2R) antiserum (Anti-I-C $^{k\,\cdot\,d}$) was also tested on B10.D2 GRF(H-2^d). It did not adsorb GRF.

region coded products (Ia antigen) complexed with small antigenic fragments with a total molecular weight of 50,000–60,000 (ref. 16). Since Ia antigens are coded for in the I region (which also contains "immune response genes") we thought that more detailed knowledge about the mechanism of action of GRF might reveal information about helper-cell induction and perhaps how immune response genes work. We have therefore studied the mechanism of binding of GRF to T cells, and report here that at least two genes, located in the I-A subregion, influence helper-cell induction. One of these genes, expressed in macrophages, controls the Ia product of GRF, and the other, expressed on T cells, controls a structure which is the receptor for GRF. This receptor does not react with anti-Ia sera and is thus distinct from conventional Ia antigen.

Though as a result of genetic experiments T cell–macrophage interaction had been mapped in the I-A subregion[10], and GRF, which is one component of this interaction, was serologically shown to be an I region product[16], it was not clear whether GRF was indeed an I-A region product or whether the receptor for GRF was I-A-controlled. Using more specific sera, directed against the H-2K^k, I-A^k, I-B^k and I-C^k regions of the *H-2* complex, it was possible to show that the gene(s) responsible for GRF lie in the I-A subregion (Table 1). The evidence for this is that only immunoadsorbents made with anti-I-A^k sera adsorbed out the activity of GRF obtained from CBA macrophages. The adsorbed activity was recovered by acid treatment of the immunoadsorbents, further demonstrating the specificity of the adsorption. An analogous mapping of the genes for GRF in the I-A subregion of other haplotypes (for example, I-A^b or I-A^d) was not possible for lack of specific appropriate anti-I-A region antisera.

We next investigated the interaction of GRF with T cells resulting in the activation of the latter to become specific helper cells. A simple explanation of this interaction would be binding of GRF by a receptor on helper-cell precursors. To test this, GRF was incubated with either T cells, B cells (10^8 cells per ml of GRF) or macrophages (6 × 10^7 cells per ml) at room temperature (RT) for 1 h. Then, after the cells had been removed by centrifugation, adsorbed GRF was tested for their helper-cell inducing capacity: T cells purified on nylon wool were used. Figure 1 shows that GRF is adsorbed by spleen cells after passage through nylon wool (containing about 1% B cells) as well as by anti-θ serum and complement-treated spleen cells (not shown), but not by macrophages. The latter point is important, making it unlikely that GRF is not synthesised by macrophages. The adsorption by anti-θ-treated spleen cells could be due to (1) residual T cells, (2) dead, but not lysed, T cells or (3) B cells. Indeed, anti-θ-treated spleen cells from which the dead cells were removed before adsorption could not adsorb out GRF activity (for example, see legend to Fig. 1). This indicates that dead cells can remove GRF, presumably because their receptors are still intact. The adsorption by T cells seems to be specific and may be related to the mechanism of GRF, action, for it is dependent on the number of T cells used for adsorption (data not shown), and because T cells incubated with GRF for 1 h at room temperature, then washed thoroughly, and cultured for 4 d were stimulated to become helper cells without further contact with GRF, macrophages or antigen (Fig. 3).

The genetics of adsorption of GRF was investigated. Allogeneic (C57BL/10=B10, H-2^b) spleen or thymus cells did not adsorb CBA GRF at a ratio of 10^8 cells per ml of GRF in contrast to syngeneic spleen cells or thymocytes (Fig. 2). Even 2 × 10^9 allogeneic cells were insufficient. To pursue the genetics of GRF adsorption further, GRF obtained from either CBA, B10.A(4R), AQR, B10 or B10.A(5R) macrophages were first adsorbed with spleen

Table 2 Genetics of the GRF-receptor on spleen cells

T cells from	Ag (μg ml^{-1})	or GRF (%)		Helper cell induction			Anti-DNP-response
				GRF adsorbed with spleen cells from	H-2 subregion shared	different	AFC ± s.e.
B10.BR	KLH 0.1			Nil	—	—	27 ± 9
		CBA GRF$_{KLH}$	10	Nil	—	—	243 ± 52
		CBA GRF$_{KLH}$	5	AQR	I-A^k,I-B^k	—	73 ± 33
		CBA GRF$_{KLH}$	15	AQR	I-A^k,I-B^k	—	57 ± 42
		B10.A(4R)GRF$_{KLH}$	10	Nil	—	—	293 ± 86
		B10.A(4R)GRF$_{KLH}$	5	CBA	K^k,I-A^k	—	30 ± 30
		B10.A(4R)GRF$_{KLH}$	15	CBA	K^k,I-A^k	—	70 ± 35
		B10.A(4R)GRF$_{KLH}$	5	AQR	I-A^k	K^k	40 ± 15
		B10.A(4R)GRF$_{KLH}$	15	AQR	I-A^k	K^k	3 ± 3
		AQR GRF$_{KLH}$	10	Nil	—	—	240 ± 30
		AQR GRF$_{KLH}$	5	B10.A(4R)	I-A^k	I-B^k	63 ± 53
		AQR GRF$_{KLH}$	15	B10.A(4R)	I-A^k	I-B^k	80 ± 55
B10	KLH 0.1			Nil	—		132 ± 108
		B10 GRF$_{KLH}$	10	Nil	—		413 ± 91
		B10 GRF$_{KLH}$	5	B10.A(4R)	I-B^b,I-C^b,S^b,D^b	K^b,I-A^b	607 ± 77
		B10.GRF$_{KLH}$	15	B10.A(4R)	I-B^b,I-C^b,S^b,D^b	K^b,I-A^b	570 ± 73
		B10.A(5R)GRF$_{KLH}$	10	Nil	—	—	670 ± 35
		B10.A(5R)GRF$_{KLH}$	5	B10.A(4R)	I-B^b	K^b,I-A^b	467 ± 61
		B10.A(5R)GRF$_{KLH}$	15	B10.A(4R)	I-B^b	K^b,I-A^b	433 ± 86
		B10.A(5R)GRF$_{KLH}$	5	B10	K^b,I-A^b,I-B^b	—	193 ± 56
		B10.A(5R)GRF$_{KLH}$	15	B10	K^b,I-A^b,I-B^b	K^b,I-A^b	123 ± 53

GRF (1ml) obtained from macrophages of various mouse strains were incubated with either A.QR, CBA, B10.A(4R) or B10 spleen cells for 1h at room temperature. The adsorbed GRF were then tested on their helper-cell-inducing capacity by incubating with either nylon wool-purified B10.BR or B10 T cells for 4 d as described in the legend to Table 1. For each combination, the H-2 subregion (of the k or b haplotype) is listed which is shared by GRF and the spleen cells used for adsorption, and in which they differ, within the left half of the H-2 complex. For example, AQR spleen cells and B10.A(4R) GRF share the I-A^k subregion but differ for the K^k subregion (lines 8 and 9). Thus the diminished activity of the B10.A(4R) GRF after adsorption with A.QR spleen cells indicates that the latter contain GRF receptors on their surface which are coded for by genes of the I-A^k and not of the K^k subregion. Five experiments of this type have been performed, for three of them nylon wool-purified T cells were used for adsorption instead of spleen cells, and other combinations were also tested with analogous results. The controls were:

B10.BR spleen cells	+ TNP-KLH (0.1 μg ml^{-1})	33 ±	27
	+ DNP–beads 3%	3030 ±	965
B10 spleen cells	+ TNP-KLH	107 ±	12
	+ DNP–beads 3%	4470 ± 1078	

Other subdivisions of the H-2I region have been described —I-E and I-J (personal communication from C. S. David and D. B. Murphy). But because these are all to the right of I-B, they do not affect the mapping studies described here.
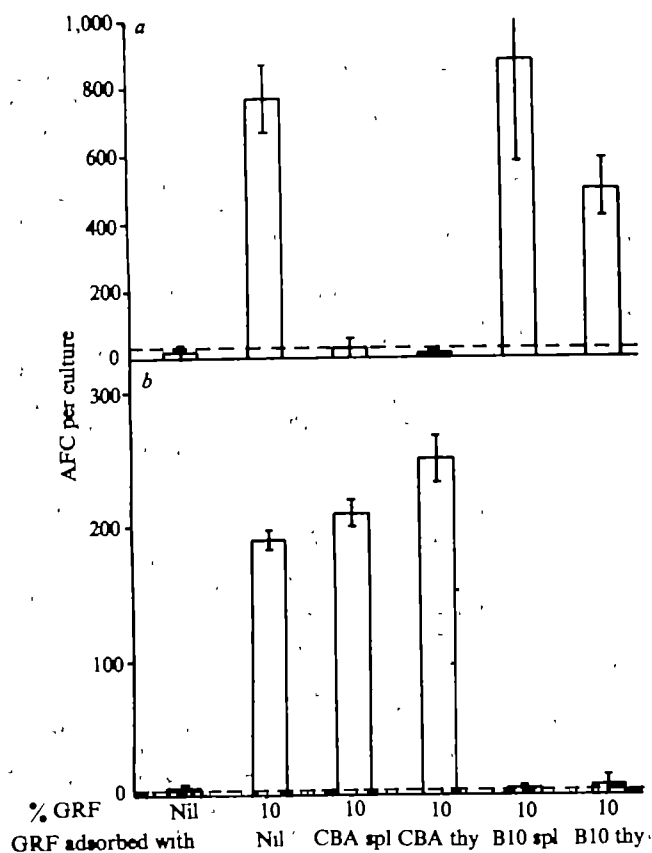
Fig. 2 Evidence for genetic control of the GRF receptor. CBA GRF$_{KLH}$ (a) or B10 GRF$_{KLH}$ (b) (1 ml) were adsorbed with either 1×10^8 CBA or B10 spleen cells (spl) or thymocytes (thy) for 1 h at room temperature. After removal of the cells by centrifugation the adsorbed GRF was tested on their helper-cell-inducing capacity by incubation with 15×10^6 nylon wool-purified CBA (a) or B10 T cells (b) in a concentration of 10% as described in the legend to Table 1. Other concentrations of GRF have also been tested, but for convenience only one concentration is shown here. As controls, unadsorbed GRF or no GRF were used. The dotted line represents the background as described in the legend to Fig. 1. Seven experiments of this type have been performed with analogous results.

cells from AQR, CBA, B10.A(4R) or B10 mice (10^6 cells per ml of GRF) and then tested for their helper-cell inducing capacity with B10.BR and B10 T cells. Table 2 shows that the adsorption of GRF by T cells is also determined by gene(s) of the *H–2* complex, located in the I-A subregion. The simplest interpretation for these results is that the receptor for GRF (which initiates the helper pathway) is controlled by genes in the I-A subregion. These results, however, do not exclude the possibility that the bulk adsorption measured here could be due to a different type of binding unrelated to the manner by which GRF stimulates helper cells. This possibility can be excluded by testing the adsorption of GRF functionally, that is whether adsorbing T cells of different strains were stimulated. Preliminary experiments with B10.A(4R) and AQR mouse strains which are identical only at the I-A$_k$ subregion of the MHC have shown that B10.A(4R) T cells are stimulated by AQR GRF. Together with the other data, this indicates that the T-cell receptor for GRF is involved in the activation of these cells, and that this T-cell receptor is controlled by the I region of the MHC.

Taken together the results indicate that at least two genes, each located in the I-A subregion, but expressed on different cells, regulate helper-cell induction. One gene codes for the Ia part of GRF, the other for the GRF receptor.

It is not likely that macrophages and T cells express both of the genes under consideration, because if so macrophages would adsorb GRF, which they do not (Fig. 1), and T cells would make GRF, which they also do not.

The possibility that the receptor for GRF, since it is controlled by the I region is an Ia antigen was tested by attempting to inhibit the adsorption of GRF by prior treatment of the T cells (which carry these receptors) with anti-Ia sera. A.TH anti-A.TL antisera, with cytotoxic titres of 1/5,000 to 1/1,000 (on spleen) and functional titres of 1/100 or more (on suppressor T cells) did not inhibit binding of GRF to the receptor, even at concentrations of 1/3, whereas the same antisera (at concentrations of 1/3, 1/30 and even of 1/1,000 in other experiments) bound GRF, if incubated together and inhibited its helper-cell-inducing capacity (Fig. 3). These results suggest that the receptor for GRF is not a classical Ia antigen.

It has been shown that the immune response to some antigens is controlled by two genes located in the *H–2* complex. Munro and Taussig[17] postulated a two-gene model for the response to the branched synthetic polypeptide (T,G)–A—L, both mapping in the I-A subregion of the *H–2* complex, with one of these genes expressed in T cells, the other in B cells. The immune response to the terpolymer L-Glu, L-Lys, L-Phe (GLØ) has also been shown to be controlled by two complementary genes in the *H–2* complex, one of which was mapped in the I-A subregion, the other somewhere to the right of the I-C subregion[18,19]. These reports deal with antigens which are clearly under immune response (*Ir*) gene control, whereas our study was performed with an antigen not known to be under *Ir* gene control. Shevach *et al.*[24] have shown in the guinea pig, however, that antisera directed against the guinea pig equivalent of the I region (for example, strain 2 anti-strain-13 sera), only inhibit responses in F$_1$ (2×13) animals controlled by the relevant I region, for example, the prolifera-
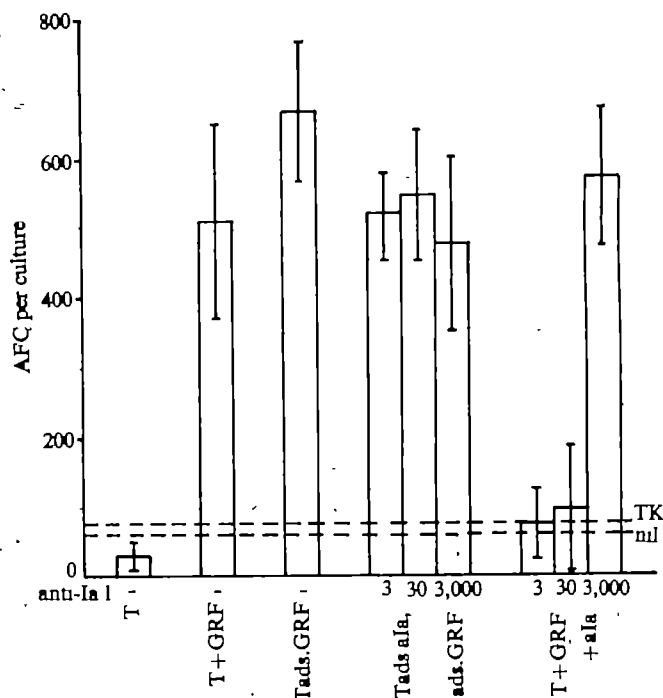
Fig. 3 The GRF receptor on T cells is not Ia antigen. 15×10^6 nylon wool-purified CBA T cells (T) were first incubated with anti-Ia serum (A.TH anti-A.TL batch no. 383) diluted 1:3, 1:30 and 1:3,000 for 3 or 4 h at room temperature. After washing, cells were incubated with CBA GRF$_{KLH}$ (0.1 ml per 15×10^6 cells) for 1 h at room temperature. After repeated washings, the T cells were incubated in Marbrook flasks for 4 d to facilitate helper-cell induction (Tads.aIa, ads. GRF). Simultaneously 15×10^6 T cells, GRF (10%) and anti-Ia^k serum in the same dilutions were incubated together for 4 d (T+aIa^k+GRF). As controls, T cells were incubated without GRF (nil), with GRF (T+GRF) or after adsorption with GRF (Tads.GRF). The procedure to measure helper activity is described in the legend to Table 1. The dotted lines represented background levels (spleen cells incubated without antigen (nil) or with TNP-KLH (TK) but without helper cells). Four experiments of this type have been performed with analogous results.

tion response to the synthetic polypeptide GT (a polymer of glutamic acid and tyrosine), an antigen the response to which is controlled by a 13-linked *Ir* gene, was inhibited by anti-13 sera. Schwartz *et al.*[25] have obtained similar results in the mouse which imply a functional relationship between *Ir* gene products and Ia antigens. Thus it is possible, even likely, that the reaction described here is a component of the function of H-2-linked *Ir* genes. At present the number of genes influencing helper-cell induction is not known, but results presented here and elsewhere indicate that at least two I region genes are involved.

The most interesting question raised by this study—the mechanism by which GRF is recognised by T cells—is still not resolved. Do helper precursor T cells have two receptors for GRF, one for the Ia product and the other for the antigen, or is there one receptor for "modified Ia" as suggested by Miller's studies[26] in delayed type hypersensitivity and in analogy to the modified H-2 involved in the killing of virus-infected cells[27] or of hapten labelled cells[28]. Our analysis does not permit an answer, but studies are in progress to resolve this issue.

We thank Dr Chella David for (A.BYxB10.HTT)F₁ anti-A.TL and (B10(4R)×129)F₁ anti-B10.A(2R) sera, Dr I. McKenzie for A.TL anti-A.TH, Dr D. Sachs for A.TL anti-A.AL serum and B10.A(4R) anti-B10.A(2R) sera, Dr D. Parker for DNP–polyacrylamide beads and Dr E. Simpson for congenic strains of mice. This investigation was supported by a grant from the Swiss National Foundation and by the Imperial Cancer Research Fund.

PETER ERB
BARBARA MEIER

Institute for Microbiology,
Petersplatz 10, CH-4003 Basel,
Switzerland

MARC FELDMANN

ICRF Tumour Immunology Unit,
Department of Zoology,
University College,
London WC1E 6BT, UK

1 Klein, J., *Biology of the Mouse Histocompatibility–2 Complex* (Springer, Berlin· Heidelberg, New York, 1975).
2 Shreffler, D. C., and David, C. S., *Adv. Immun.*, 20, 125–195 (1975).
3 Benacerraf, B., and Katz, D. H., *Immunogenetics and Immunodeficiency* (edit. by Benacerraf, B.), 117 (Medical and Technical Publishing Co., London, 1975).
4 Miller, J. F. A. P., Vadas, M. A., Whitelaw, A., and Gamble, J., *Proc. natn. Acad. Sci. U.S.A.*, 72, 5095–5098 (1975).
5 McDevitt, H. O., and Benacerraf, B., *Adv. Immun.*, 11, 31–74 (1969).
6 Klein, J., Hauptfeld, M., and Hauptfeld, V., *Immunogenetics*, 1, 45 (1974).
7 Shreffler, D. C., David, C. S., Gotze, D., Klein, K., McDevitt, H. O., and Sachs, D., *Immunogenetics*, 1, 189 (1974).
8 Benacerraf, B., and McDevitt, H. O., *Science*, 175, 273–279 (1972).
9 Munro, A. J., Campbell, M. J., Taussig, M. J., Williams, H., and Lawson, Y., *J. exp. Med.*, 140, 1579–1587 (1974).
10 Erb, P., and Feldmann, M., *J. exp. Med.*, 142, 460–472 (1975).
11 Kontiainen, S., and Feldmann, M., *Nature new Biol.*, 245, 285–286 (1973).
12 Erb, P., and Feldmann, M., *Nature*, 254, 352–354 (1975).
13 Erb, P., and Feldmann, M., *Cell. Immun.*, 19, 356–367 (1975).
14 Shevach, E. M., and Rosenthal, A. S., *J. Exp. Med.*, 138, 1213–1229 (1973).
15 Erb, P., and Feldmann, M., *Eur. J. Immun.*, 5, 759–766 (1975).
16 Erb, P., Feldmann, M., and Hogg, N., *Eur. J. Immun.*, 6, 365–372 (1976).
17 Munro, A. J., and Taussig, M. J., *Nature*, 256, 103–106 (1975).
18 Dorf, M. E., Stimpfling, J. H., and Benacerraf, B., *J. exp. Med.*, 141, 1459–1463 (1975).
19 Merryman, C. F., Maurer, P. H., and Stimpfling, J. H., *Immunogenetics.*, 1, 549–559 (1975).
20 Merryman, C. F., Maurer, P. H., and Stimpfling, J. H., *Immunogenetics* (in the press).
21 Dorf, M. E., and Benacerraf, B., *Proc. natn. Acad. Sci. U.S.A.*, 72, 3671–3675 (1975).
22 Schwartz, R., Dorf, M. E., Benacerraf, B., and Paul, W. E., *J. exp. Med.*, 143, 897–905 (1976).
23 Katz, D. H., Dorf, M. E., and Benacerraf, B., *J. exp. Med.*, 143, 906–918 (1976).
24 Shevach, E. M., Paul, W. E., and Green, I., *J. exp. Med.*, 136, 1207–1221 (1972).
25 Schwartz, R. H., David, C. S., Sachs, D. H., and Paul, W. E., *J. Immun.* 117, 531–540 (1976).
26 Miller, J. F. A. P., Vadas, M. A., Whitelaw, A., and Gamble, J., *Proc. natn. Acad. Sci. U.S.A.* 73,, 2486–2490 (1976).
27 Zinkernagel, R. M., and Doherty, P C , *J. exp. Med.*, 141, 1427–1436 (1975).
28 Shearer, G. M., Rehn, T. G., and Garbarino, C. A., *J. exp. Med.*, 141, 1348–1364 (1975).
29 Von Boehmer, H., and Shortman, K., *J. Immun. Med.*, 2, 293–301 (1973).

Liposome-induced morphological differentiation of murine neuroblastoma

LIQUID crystals made of lipids (liposomes), initially devised as models of biological membranes[1–4], have more recently

been used as artificial carriers because of their ability to bind and fuse with biological membranes and discharge their soluble content into the cytoplasm[5–7]. Preliminary studies in our laboratory have shown that lisopsomes loaded with the nerve growth factor (NGF) induce morphological differentiation in a murine neuroblastoma cell line which does not exhibit morphological changes on addition of NGF to the culture medium. Control experiments with liposomes in absence of NGF however, also resulted in process formation in neuroblastoma (NB) cells. Our attention was therefore focused on the effect of the carrier itself, rather than on the agent carried.

A neuroblastoma cell clone (NB41A₃) was obtained from Dr Augusti-Tocco[8] and maintained in suspension in F10 medium (Gibco), supplemented with 15% horse serum and 2.5% foetal calf serum (EUROBIO, Paris), penicillin G (200 U ml⁻¹), and streptomycin sulphate (200 μg ml⁻¹) in an humidified incubator conditioned with 5% CO₂ and 95% air, at 37 °C. Egg yolk phosphatidyl choline (PC), phosphatidyl serine (PS), phosphatidyl ethanolamine (PE), phosphatidic acid (PA) and sphingomyelin (SM) were purchased from Lipid Products, South Nutfield, Surrey. N⁶,O²-dibutyryl cyclic AMP, 5-bromodeoxyuridine (BUdR) and papaverine were obtained from Sigma, dimethylsulphoxide (DMSO) was from Merck.

The first set of experiments was directed to find the optimum conditions (liposome composition, time of incubation, temperature) for maximum binding. This process was measured by the amount of ¹⁴C–PC mixed with the phospholipids forming the liposomes, which was not removed by three consecutive washings of the cells after incubation. This measure of 'tightly bound' liposomes does not, however, indicate whether binding is followed by fusion and/or interiorisation within the cell cytoplasm. These experiments indicated that a mixture of sonicated PC–PS 10 : 1.5 (w/w) liposomes progressively attached to, and/or penetrated the cell membrane during a 24-h incubation period at 37 °C. At the end of this time, 0.2–0.4% of the total liposomes present in culture was tightly bound to the cells. Addition of cholesterol to this lipid suspension did not significantly change the pattern of binding. Changing the net charge of the liposome from negative to positive by introducing stearylamine severely reduced the viability of NB cells.

The marked morphological changes (process formation) induced by liposomes are shown in Fig. 1*a* and *b*, which compares cultures of the same NB cell line in the absence or presence of this lipid suspension. The fine structure of the liposome-induced processes (Fig. 1*c*) shows an abundance of microtubules (MT) lying in parallel arrays. In addition to MTs, other fibrillar structures tentatively identified as microfilaments (MF) are also seen in most liposome-treated cultures. The number of cells present in cultures containing liposomes is 40% lower after 24 h and 60% lower after 48 h, than in control cultures. During the same period, the TCA precipitable ³H-methylthymidine is reduced to 30% of controls when expressed as counts per mg of total protein. These findings suggest that liposome-induced morphological differentiation is accompanied by a concomitant inhibition of cell division, although we cannot yet decide whether the two phenomena are correlated or independent. Autoradiographic studies show heavily labelled NB cells following incubation with ¹⁴C liposomes for after only 3 h of incubation (Fig. 2). Whether the radioactivity is confined to the cell surface, or whether it is also within the cytoplasm, is not known.

Several substances have been shown to induce morphological or biochemical maturation of NB cells[9–13]. In our NB system, however, only dibutyryl cyclic AMP induced process formation, although not to the same extent as liposomes: other agents such as papaverine, 5-bromodeoxyuridine and dimethylsulphoxide, had little or no effect. It is worth mentioning that PS liposomes injected *in vivo*,
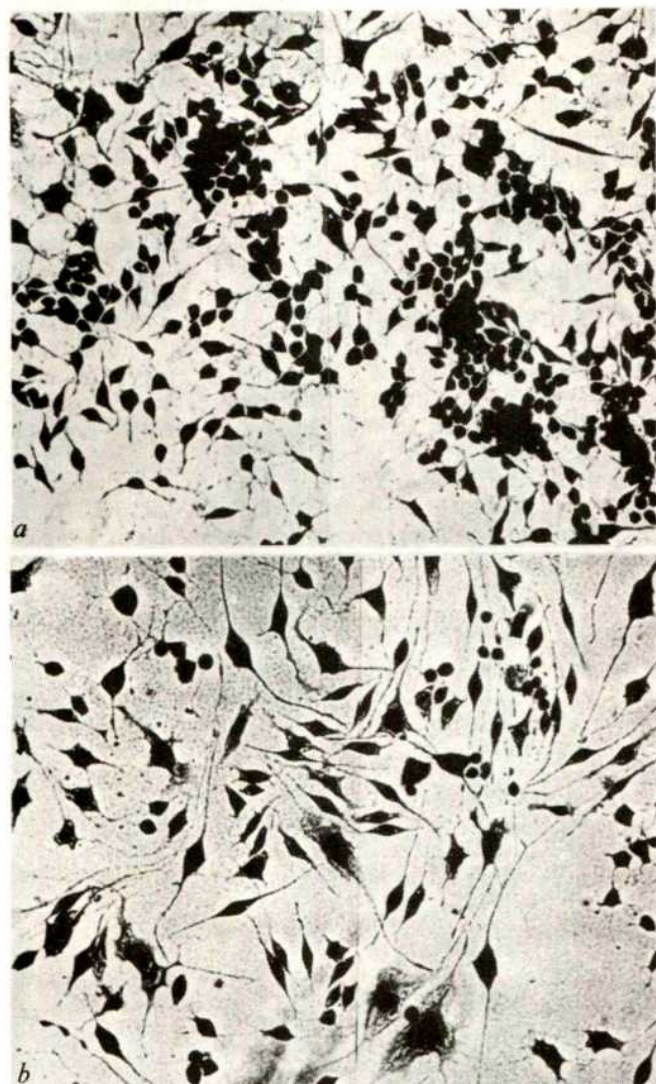
Fig. 1 Effect of liposomes on the morphology of neuroblastoma cells. NB cells (NB41A$_3$) were plated in a Falcon dish containing coverslips. After 3 h, 100 μl (2.7 mg of total phospholipids) of the suspension of liposomes was added to a series of cultures (*b*, *c*) or the same volume of Gey's balanced salt solution was added to control cultures (*a*). After 3 d, the cultures were fixed either for light microscope observation (*a*, *b*) or for electron microscopic studies (*c*). *a*, *b*, Microphotographs taken using phase contrast were fixed with Dubosc Brasil and stained with haematoxylin. *c*, Electron micrography of a portion of neurite from the same culture as shown in (*b*). Cells were fixed in 3% glutaraldehyde in 0.1 M phosphate buffer with Ca^{2+} added in the same buffer and post-fixed in 1% OsO$_4$, dehydrated and embedded in Epon 812. Photo was taken with Philips E.M. 300. Magnifications: *a* and *b*, ×174; *c* ×2,079. Liposomes were prepared as described previously[1-3]. Generally, 20 mg lecithin from egg yolk was mixed with 3.0 mg phosphatidylserine in 1.0 ml of chloroform–methanol 2:1 (v/v) and dried under a vacuum pump. Gey's balanced salt solution (1 ml) was added and stirred vigorously for 5 min in a vortex mixer. Sonication was in a constant flow of nitrogen gas for 25 min. Liposomes obtained in this way are round, with a diameter of about 120–200 Å measured after ammonium molybdate staining, in the electron microscope.

have been shown to increase brain glucose and decrease catecholamine content of the adrenal medulla[14,15].

Our findings raise the question as to the mechanism by which pure lipids, which are normal constituents of the cell membrane, can induce morphological maturation of a neoplastic cell line. Using ^{14}C–PC-loaded liposomes incubated at a concentration of 1.0 μmol^{-1} with 4×10^6 cells, we have found that 2–4 nmol of liposomes are bound to, or have penetrated, the cells. This value corresponds to an average of 3×10^8 phospholipid molecules per neuroblastoma cell, that is, (assuming that a sonicated liposome is formed by 2×10^3 phospholipid molecules[2]), 1.5×10^5 liposomes bound per cell. A liposome bound to the membrane of a living cell may undergo endocytosis, or it may fuse with the membrane itself and become part of its constituents. If the fate of phospholipids is mainly to become constitutent parts of the cytoplasmic membrane, their contribution to the

Fig. 2 Autoradiograph of NB41A$_3$ treated with radioactive liposomes. PC-PS liposomes were prepared as described in Fig. 1 in the presence of 1 μCi ml^{-1} ^{14}C–PC (New England Nuclear, specific activity 1.765 Ci mmol^{-1} P). After incubation for 3 h, cells (which were grown on coverslips in a Falcon dish) were fixed in 3% glutaraldehyde in 0.1 M phosphate buffer *p*H 7.2. After 1 h in fixative, cultures were washed in the same buffer, and stained in haematoxylin. The coverslips were air dried and coated with Ilford L4 emulsion. After 50 d of exposure at room temperature, the autoradiograms were developed with Microdol at 21 °C for 6 min. Microphotographs were taken with: *a*, transmission light; *b*, interference differential (Zeiss–Nomarski) microscope. Magnifications for *a* and *b* ×824.
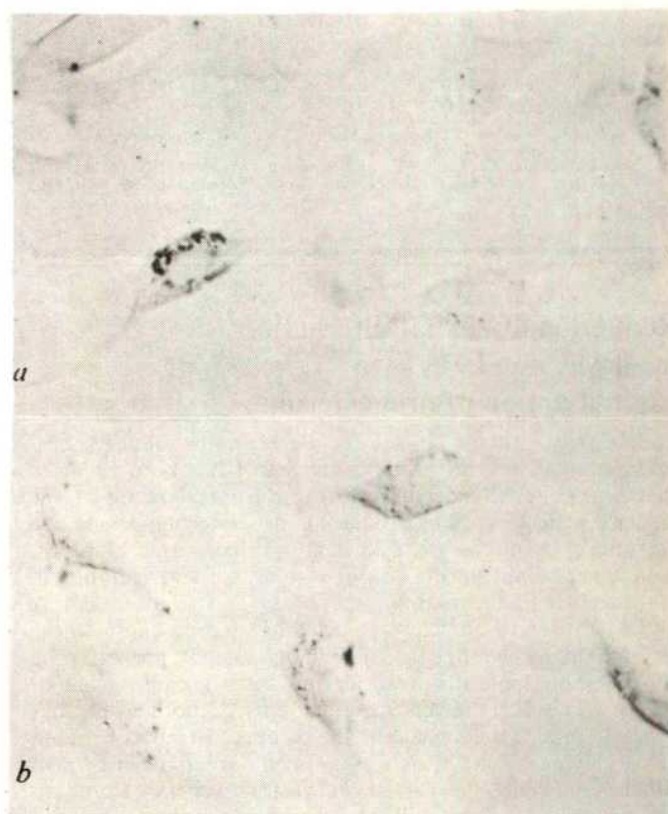
entire surface of each NB cell of diameter 15 μm could be between 5 and 20%, a not insignificant increase for the functional properties of a membrane. Considering that PS molecules are negatively charged and exhibit particular binding capacity for Ca^{2+} and monovalent cations, changes in the net charge of the membrane can be expected following binding and/or fusion of the phospholipids forming the liposomes. The considerable amount of bound phospholipids might also influence the permeability of these cells and effect the organisation of the fibrillar proteins (microtubules, microfilaments) which are essential in axonal growth and elongation.

It remains to be ascertained to what extent *in vitro* process formation in NB cells is a true morphological sign of neural differentiation, since this effect can be induced not only by specific substances[10-13] but also by noxious agents or treatments[11,12]. The use of liposomes may contribute to this controversial problem, since their action seems to involve a mechanism which is not dependent on a specific signal, nor do they cause a detrimental effect on the responsive system.

We thank Dr R. Levi-Montalcini for discussions during the preparation of the manuscript and Dr Augusti-Tocco for supply of NB41A3 cells.

J. S. CHEN
A. DEL FA
A. DI LUZIO
P. CALISSANO

*Laboratorio Biologia Cellulare CNR,
Via G. Romagnosi, 18A,
Rome Italy*

1 Bangham, A. D., in *Permeability and Function of Biological Membranes* (edit. by Bolis, L., *et al.*), 195–205 (North-Holland, Amsterdam, 1970).
2 Johnson, S. M., Bangham, A. D., Hill, M. W., and Korn, E. D., *Biochim. biophys. Acta*, 233, 820–826 (1971).
3 Calissano, P., and Bangham, A. D., *Biochem. biophys. Res. Commun.*, 43, 504 (1971).
4 Calissano, P., Alema, S., and Fasella, P., *Biochemistry*, 13, 4553–4559 (1974).
5 Martin, F., and MacDonald, R., *Nature*, 252, 161–163 (1974).
6 Paphadjopoulos, D., Mayhew, E., Poste, G., Smith, S., and Vail, W. J., *Nature*, 252, 163–165 (1974).
7 Pagano, R. E., Huang, L., and Wey, C., *Nature*, 252, 166–167 (1974).
8 Augusti-Tocco, G., Sato, G., Claude, P., and Potter, D., in *Symp. Int. Soc. Cell Biol.*, 9 (edit. by Padycula, H. A.), 109–120 (Academic, New York, 1970).
9 Seeds, N. W., Gilman, A. G., Amano, T., and Niremberg, M. W., *Proc. natn. Acad. Sci. U.S.A.*, 66, 160–167 (1970).
10 Prasad, K. N., and Hsie, A. W., *Nature new Biol.*, 233, 141–142 (1971).
11 Schubert, D., and Jacob, F., *Proc. natn. Acad. Sci. U.S.A.*, 67, 247–254 (1970).
12 Kimhi, Y., Palfrey, C., Spector, I., Barak, Y., and Littauer, U. Z., *Proc. natn. Acad. Sci. U.S.A.*, 73, 462–466 (1973).
13 Monard, D., Solomon, F., Rentsch, M., and Gysin, R., *Proc. natn. Acad. Sci. U.S.A.*, 70, 1894–1897 (1973).
14 Bruni, A., Toffano, G., Leon, A., and Boarato, E., *Nature*, 260, 331–333 (1976).
15 Amaducci, L., Mantovani, P., and Pepeu, G., *Br. J. Pharmac.*, 56, 739–780 (1976).

Reduction of sexual interaction in rhesus monkeys by a vaginal action of progesterone

PROGESTERONE acting in the female seems to reduce sexual interaction in several primate species. It has been reported that sexual activity declines during the luteal phase of the menstrual cycle (when circulating progesterone levels are maximal), in monkeys[1-3], lowland gorillas[4] and humans[5], and that administering progesterone to ovariectomised, oestrogen-treated rhesus monkeys has a similar effect[6]. A major unresolved question is how progesterone causes these changes in behaviour. One possibility is that progesterone acts on the female's central nervous system, causing her to accept or solicit fewer male mounts. Another is that progesterone somehow alters the vagina thereby changing non-behavioural cues (such as a smell[7] or tactile qualities) which contribute to her sexual attractiveness. Here we present evidence favouring the second mechanism, since the

reduction in sexual interaction caused by systemic administration of a physiological dose of progesterone to female monkeys could be reproduced by instilling very small amounts of this hormone directly into the vagina.

Ovariectomised female and intact male adult rhesus monkeys (*M. mulatta*) were used. The females were implanted subcutaneously with silastic tubing (Dow Corning, 4 cm long, 3.35 mm inner and 4.65 mm outer diameters) filled with oestradiol-17β, which results in plasma levels of oestradiol similar to those measured in the follicular phase of the menstrual cycle[8]. Animals were housed separately between observations. Sexual interaction between heterosexual pairs of monkeys selected because they copulated readily without displaying aggression was studied from behind a one-way mirror[9]. In each experiment the same pairs were tested daily in a constant order until the male ejaculated, or for a maximum of 30 min. Here we consider only the following behavioural measures: (1) for the male, mounting rate (number of mounts per min) and the percentage of tests with ejaculation; and (2) for the female, acceptance ratio (the percentage of male mounting attempts accepted) and the number of sexual invitations[10] offered to the male per observation. Plasma progesterone was estimated by radioimmunoassay, using at least two dilutions of plasma[11].

In experiment 1 six females were given daily intramuscular injections of 10 mg progesterone in arachis oil, which after 20 d resulted in plasma levels of progesterone (median: 10.74 ng ml⁻¹, range: 5.16–14.52) within the
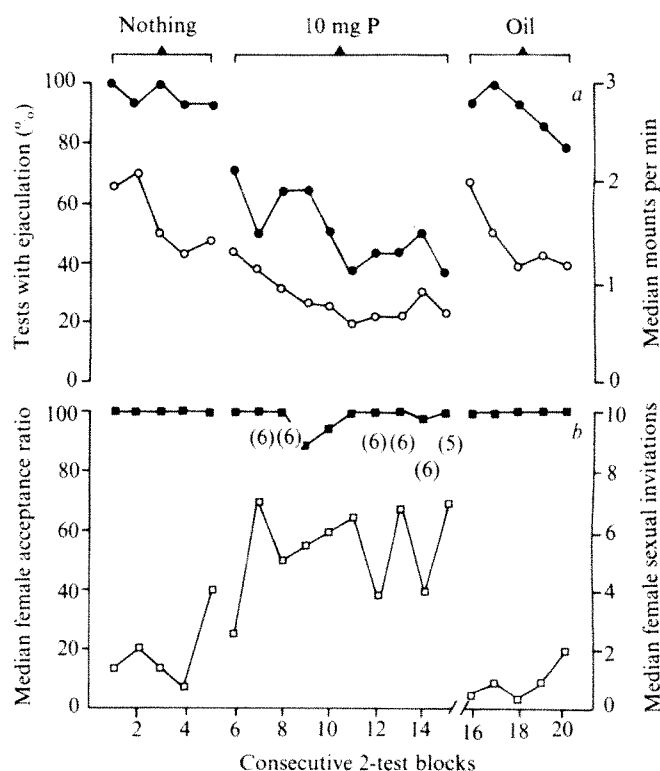
Fig. 1 Effect of 10 mg d⁻¹ progesterone given intramuscularly to the female on sexual interaction in pairs of rhesus monkeys (experiment 1). Five females were tested with one of three males, and another female with two different males. An interval of 32 d, during which no treatment or tests were administered, intervened between test blocks 15 and 16. Two parameters of the males' sexual behaviour are shown in *a*: ●, % tests with ejaculation; ○, median mounts min⁻¹; and two of the females' in *b*: ■, median acceptance ratio; □, median sexual invitations. Data points are based on seven pairs. Exceptions, which are indicated in parentheses, represent tests in which no acceptance ratio could be calculated for some pairs. The statistical comparisons were made with the initial control condition using 2-tailed Wilcoxon tests, provided an initial Friedman analysis of variance was significant ($P < 0.05$).
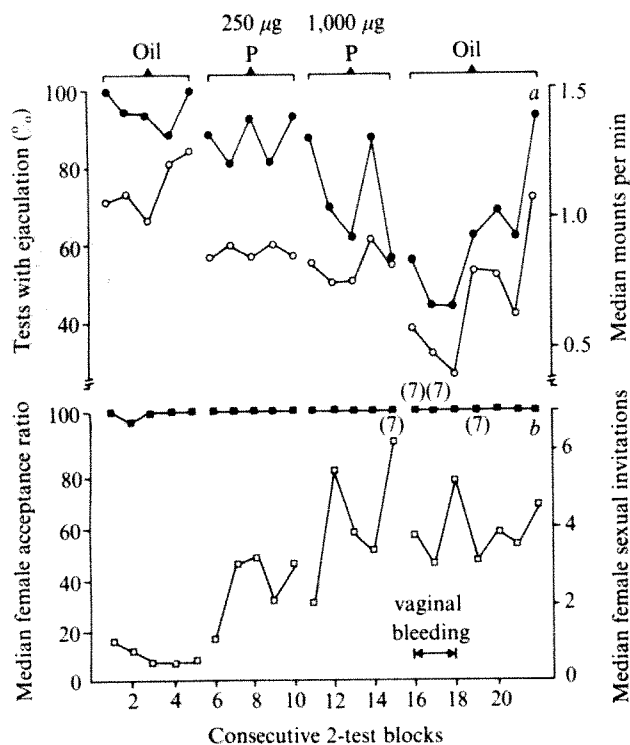
Fig. 2 Effect of low doses of progesterone intravaginally on sexual interaction (experiment 2). Six females were tested with one of six males, and another female with two different males. After each day's behavioural testing, arachis oil solutions were administered by inserting a catheter adaptor attached to a 1–ml syringe into the vagina until it touched the cervix, when it was withdrawn slightly and the injection made. Data points are based on eight pairs, exceptions being indicated in parentheses. Behavioural parameters (*a*, male; *b*, female), statistical analysis and symbols as in Fig. 1.

range measured during the luteal phase of the menstrual cycle[8]. This treatment caused significant reductions in the males mounting rate and ejaculation ($P<0.02$), whereas the females continued to accept virtually all the mounting attempts and actually displayed a threefold increase in sexual invitations ($P<0.02$) (Fig. 1). Thus, physiological increases in circulating progesterone made the females significantly less attractive to male partners, without altering neural mechanisms so as to depress receptivity. Our results contrast with the reported[6] reduction in receptivity in some female monkeys which occurred following administration of a pharmacological dose of progesterone (25 mg d^{-1}). In a second experiment, we investigated the possibility that

Table 1 Plasma progesterone in seven ovariectomised, oestradiol-treated female rhesus monkeys following intravaginal and intra-muscular administration of progesterone

Site and amount of progesterone administered in 0.1 ml arachis oil	Treatment duration (d)	Plasma progesterone (ng–ml) (Median)	(Range)
A Vagina			
0	10	0.12	<0.10–0.41
250 µg	10	0.29	<0.10–0.43
1,000 µg	10	0.30*	0.15–0.52
0	14	0.15	<0.10–0.24
B Thigh muscle			
0	10	0.20	<0.10–0.51
250 µg	10	0.59*	0.28–0.83
1,000 µg	10	0.81†	0.63–1.32

Within each experiment, comparisons were made with the initial 0 condition using one-tailed Wilcoxon tests. Blood samples were collected 24 h after the last injection in each treatment block. *$P<0.05$, †$P<0.01$.

progesterone affected sexual interaction by acting directly on the vagina.

In experiment 2, instilling progesterone directly into the vagina duplicated the behavioural effects obtained with systemic administration in experiment 1. Figure 2 shows that giving 250 µg d^{-1} progesterone intravaginally caused a significant decline in mounting rate and a significant increase in sexual invitations by the females ($P<0.02$). The incidence of ejaculation declined when the dose of progesterone was raised to 1,000 µg d^{-1}; however, masculine sexual performance returned to initial control levels after intravaginal progesterone treatment was stopped. Females' acceptance ratios remained high throughout the experiment. Intravaginal administration of 250 µg progesterone had no significant effect on plasma levels, although treatment with 1,000 µg caused a slight, though significant increase (Table 1*A*). These findings suggest that progesterone altered sexual interaction primarily by acting on the vagina. The increased display of sexual invitations by females which was observed did not, therefore, result from the direct action of progesterone on neural tissues controlling this behaviour. Instead, it may reflect an attempt by females to compensate behaviourally for a progesterone-induced reduction in sexual attractiveness.

Fig. 3 Effect of low doses of progesterone intramuscularly on sexual interaction (experiment 3). Six females were tested with one of five males. Data points are based on six pairs, exceptions being indicted in parentheses. Behavioural parameters (*a*, male; *b*, female), statistical analysis and symbols as in Fig. 1.

These conclusions were supported by the results of experiment 3, in which the same low doses of progesterone were administered intramuscularly instead of intravaginally. Plasma levels of progesterone increased significantly when 250 µg d^{-1} progesterone was given intramuscularly (Table 1*B*); however, this treatment had no significant effect on sexual interaction (Fig. 3). Giving 1,000 µg progesterone raised plasma progesterone levels even more and caused a significant decline in males' mounting rate and ejaculation ($P<0.05$). The females' receptive behaviour was not affected, although in this experiment the increase in their sexual invitations induced by progesterone was not significant. The fact that progesterone in the lower dosage had to be instilled directly into the vagina (experiment 2) to

affect sexual interaction demonstrates that the behavioural effects of this steroid result from its action in this tissue.

Oestradiol enhances sexual attractiveness of female monkeys by affecting the vagina[11]. It may act by stimulating the emission of olfactory attractant cues[7] or by altering the tactile qualities of this tissue. Progesterone could reduce the sexual attractiveness of the females by blocking these effects of oestradiol in the vagina. A comparable anti-oestrogenic action of progesterone has been demonstrated in the oviducts of rhesus monkeys[11]. Alternatively, progesterone could act synergistically with oestradiol to stimulate the production of a vaginal olfactory cue which actively reduces female sexual attractiveness. Either mechanism could account for the reduction in sexual attractiveness which has been reported in different primate species[14] (including humans[9]) at times when the female's plasma progesterone levels are elevated.

M. J. BAUM*
B. J. EVERITT
J. HERBERT
E. B. KEVERNE

*Department of Anatomy,
University of Cambridge,
Cambridge CB2 3DY, UK*

W. J. DE GREEF

*Department of Endocrinology,
  Growth and Reproduction,
Erasmus University, Rotterdam,
The Netherlands*

1 Michael, R. P., and Zumpe, D., *J. Reprod. Fert*, 21, 199–201 (1970).
2 Bullock, D. W., Para, C. A., and Goy, R. W, *J Reprod Fert.*, 31, 225–236 (1972).
3 Saayman, G. S., *Folia Primat.*, 12, 81–110 (1970).
4 Nadler, R. D., *Science*, 189, 813–814 (1975).
5 Udry, J. R., Morris, N M., and Waller, L, *Arch sex Behav.*, 2, 205–214 (1973).
6 Michael, R. P., Saayman, G. S, and Zumpe, D, *J. Endocr*, 41, 421–431 (1968).
7 Michael, R P., Keverne, E. B., and Bonsall, R. W, *Science*, 172, 964–966 (1971).
8 Knobil, E., *Rec. Prog. Hormone Res.*, 30, 1–46 (1974).
9 Everitt, B. J., Herbert, J., and Hamer, J. D, *Physiol. Behav*, 8, 409–415 (1972).
10 Michael, R. P., and Zumpe, D., *Behaviour*, 36, 168–186 (1970).
11 de Jong, F H., Baird, D T., and van der Molen, H J, *Acta Endocr*, 77, 575–587 (1974).
12 Keverne, E. B., in *Advances in the Study of Behaviour*, 7 (edit. by J. S Rosenblatt), (Academic, in the press).
13 Brenner, R. M., Resko, J. A, and West, N. B, *Endocrinology*, 95, 1094–1104 (1974).
14 Eaton, G. G., in *Primate Reproductive Behaviour* (edit. by Phoenix, C. H), 20–35 (Karger, Basel, 1973).

Hypertensive effect of 17α, 20α-dihydroxyprogesterone and 17α-hydroxyprogesterone in the sheep

ANIMAL experimental models of high blood pressure as caused, for example, by renal artery constriction or administration of mineralocorticoid hormones and salt, have yielded basic data on physiological processes which have helped in the diagnosis and treatment of human hypertension. The discovery that ACTH administration causes a reproducible hypertension of rapid onset in the sheep provided a new model for analysis[1,2]. Certain clinical observations have indicated that there may be as yet unidentified steroids which are causally involved in some hypertensive states: for example, instances of high blood pressure in children and adults which were corrected by dexamethasone treatment[3,4], and reversal of high blood pressure in the low renin group of essential hypertensives by inhibitors of steroidogenesis or by spironolactone have been described[4-6]. We report here the hypertensive effect of 17α, 20α-dihydroxyprogesterone

(17α,20α-diOHP; 17α,20α-dihydroxy-4-pregnene-3-one) and 17α-hydroxyprogesterone (17α-OHP) in sheep made hypertensive with ACTH. These two progesterones were identified in the adrenal venous effluent from ACTH-hypertensive sheep and were shown to be secreted in greatly increased amounts.

Animal experimental procedures have been described in detail previously[1,2]. The animals were habituated over several weeks of training to blood pressure measurement by the same observers, and the daily mean was based on 6–8 readings during morning and afternoon. In the absence of a recognisable disturbance in the laboratory, variation in the trained animal during the day was small—less than 5 mmHg. After a control 3–5-d period, intravenous steroid infusions were carried out over 5 d using the same measurement protocol as in the ACTH and the other steroid infusion experiments[1,2].

ACTH administration (80 IU d^{-1}) to the intact, but not the adrenalectomised sheep, results in a significant elevation of arterial blood pressure within 24 h and this is sustained over 5–10 d of ACTH treatment[1] (Fig. 1, group e). Hypokalaemia, an increase in plasma sodium, and increased water intake and urine output were also observed[1]. Changes in blood pressure were not related to changes in external body Na status or in body weight. Intravenous infusion of cortisol, corticosterone, deoxycorticosterone (DOC), aldosterone and 11-deoxycortisol either singly, or in combination at rates to give blood levels similar to those found during ACTH administration (cortisol, 5 mg h^{-1}; corticosterone, 0.5 mg h^{-1}; 11-deoxycortisol, 1 mg h^{-1}; DOC, 25 μg h^{-1}; aldosterone, 3 μg h^{-1}), failed to reproduce the elevation in blood pressure[2] (Fig. 1, group a). The combined steroid infusion, however, did reproduce all other metabolic responses seen in the ACTH experiments[1,2]. In addition, the combined steroids together with 18-hydroxydeoxycorticosterone (18-OHDOC, 100 μg h^{-1}) did increase blood pressure on the fifth day[2] (Fig. 1, group b). This has been confirmed in an additional six experiments.
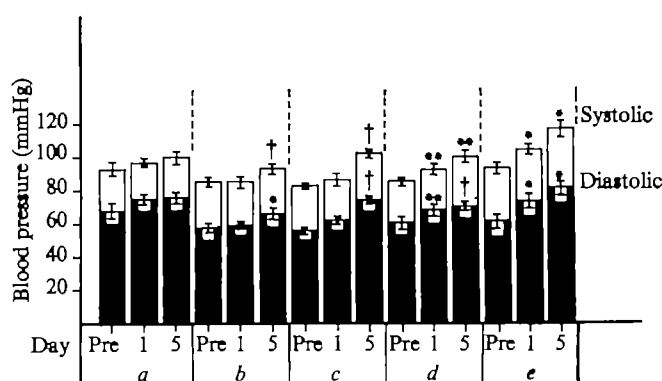
Fig. 1 Mean systolic and diastolic blood pressure (±s.e.m.) before (pre) and after 1 and 5 d of the combined steroid infusion (group a); combined steroid plus 18-OHDOC (group b); combined steroid plus 17α-OHP (group c); combined steroid plus 17α,20α-diOHP (group d), and ACTH (group e). Statistical comparisons with control (pre) observations have been made using Students' t test for paired observations. *P < 0.05; **P < 0.01; †P < 0.005.

The animals in the first group (n = 5) were given the combined steroid infusion together with 17α-OHP (Steraloids Inc.) at a rate of 1 mg h^{-1} for 5 d. This rate approximated the rate of adrenal secretion of the hormone measured during ACTH-induced hypertension in conscious trained sheep with an adrenal transplant[7], or with an adrenal-jugular bypass operation[8]. The administration rate was thus in the physiological range of adrenal functional capacity to secrete the hormone. Systolic blood pressure increased from 82 ± 1 (mean ± s.e.) to 94 ± 2 and 102 ± 2 mmHg on the fourth (P < 0.05) and fifth day (P < 0.005), respectively (Fig. 1, group d). Diastolic blood pressure showed a significant increase 56 ± 1–61 ± 1 mmHg on

the second day ($P < 0.005$). Neither systolic nor diastolic blood pressure was elevated within the first 24 h.

Animals in the second group ($n = 6$) were given a combined steroid infusion together with 17α,20α-diOHP (Steraloids Inc.) at 500 μg h^{-1}. As in the first group, this rate approximated the physiological output determined during ACTH-induced hypertension and both systolic and diastolyic blood pressure were raised significantly within 24 h. Systolic pressure rose from a control of 85 ± 2 to 92 ± 3 mmHg ($P < 0.01$) (Fig. 1, group c); on the final day of infusion, it was 100 ± 3 mmHg ($P < 0.01$). Diastolic blood pressure rose from a control of 60 ± 3 to 68 ± 3 mmHg ($P < 0.01$) and was significantly elevated at 70 ± 2 mmHg on the fifth infusion day ($P < 0.005$).

Preliminary experiments in two animals indicated that neither 17α-OHP nor 17α, 20α-diOHP, when infused alone at the same rate as described above, caused any increase in blood pressure over a 5-d period. In a further two experiments where they were given together with the combined steroid infusion, the blood pressure response was similar to that seen in the second group of experiments.

In the study with combined steroid plus 17α,20α-diOHP, urinary sodium response involved a rather more sustained initial sodium retention than with ACTH, although blood pressure effects were similar. All other metabolic responses were similar to those in response to ACTH. The more sustained urinary Na retention occurred both with 17α-OHP with combined steroid, and when 17α,20α-diOHP and 17α-OHP were infused together with the combined steroid mixture. When either of these steroids was infused alone, without the concurrent combined steroid infusion, they were without effect on urinary sodium excretion.

There is, however, evidence that urinary sodium retention is not the primary mechanism of ACTH hypertension. Sodium depletion of 15–20% of the exchangeable sodium pool by acute parotid duct cannulation before the start of 5 d of ACTH treatment, reduced urinary Na excretion to basal levels. In six animals this did not alter the time course of the blood pressure response to ACTH[9]. Sodium depletion was also carried out in seven animals after the 5-d ACTH treatment, with ACTH being continued over a further 2 d. The blood pressure, though reduced by depletion was still significantly elevated at the end of the sodium depletion period[9]. In five experiments where ACTH was given 24 h after bilateral nephrectomy, there was no difference in blood pressure response from that of intact animals. Although these data support the theory that urinary sodium retention is not the cause of the ACTH hypertension, we cannot formally exclude the possibility that it may be influential in the hypertensive effect of 17α-OHP and 17α,20α-diOHP plus combined steroid infusions.

Hypertension associated with excessive endogenous steroid production or administration of exogenous steroids, has been widely reported clinically and in experimental studies. Whereas the cause of increased blood pressure in some of these studies is well known, with others, the cause(s) are not clear. Historically, two classes of steroids—mineralocorticoids such as aldosterone, DOC and corticosterone, and glucocorticoids such as cortisol and 11-deoxycortisol—have been believed to be able to cause hypertension. The exact mechanism by which these types of hypertension are produced and the precise relationship between these steroids, sodium status and blood pressure are still largely undefined. Other naturally occurring steroids—for example, 18-OHDOC and 16-βOH-dehydroepiandrosterone—have been investigated for their hypertensive activities[4,10]; the former increases blood pressure in the dog[11] and sheep[2] but any hypertensive action of the latter requires confirmation. In the present study, two steroids previously unrecognised as having an hypertensive action were shown to increase blood pressure in sheep when given at high physiological dosages concurrent with, and conditional on, elevated levels of a number of other adrenocortical hormones.

The mechanism by which these two steroids influence blood

pressure is not yet clear. The fact that alone, they do not increase blood pressure, nor do they reduce the salivary Na–K ratio following ipsilateral carotid artery infusion in the Na-deficient adrenalectomised sheep, suggests they are not classical mineralocorticoids. Because doubling the cortisol component of the combined steroid infusion does not reproduce the blood pressure increase of ACTH or addition of 17α,20α-diOHP, it is unlikely that they simply add to an existing glucocorticoid-type effect. If they are not glucocorticoids, then recognition of a class of steroids which can increase blood pressure, but which may not act as classical mineralocorticoids, could have important implications for future clinical research and experimental models of hypertension.

J. P. COGHLAN
D. A. DENTON
J. S. K. FAN
J. G. MCDOUGALL
B. A. SCOGGINS

Howard Florey Institute of Experimental
*   Physiology and Medicine,*
University of Melbourne,
Parkville 3052, Australia

1 Scoggins, B. A., *et al.*, *Am. J. Physiol.*, **226**, 198 (1974).
2 Fan, J. S. K., *et al.*, *Am. J. Physiol.*, **228**, 1695 (1975).
3 New, M. I., Saenger, P. H., Peterson, R. E., and Ulick, S., *Bull. N. Y. Acad. Med.*, **51**, 1179 (1975).
4 Melby, J. C., Dale, S. L., Grekin, R. J., Gaunt, R., and Wilson, T. E., *Rec. Prog. Hormone Res.*, **28**, 287 (1972).
5 Woods, J. W., Liddle, G. W., Stant, E. G., Michelakis, A. M., and Brill, A. G., *Arch. intern. Med.*, **123**, 366 (1969).
6 Gunnells, J. C., *et al.*, *Ann. intern. Med.*, **73**, 901 (1970).
7 Wright, R. D., *et al.*, *Austr. J. exp. Biol. med. Sci.*, **50**, 873 (1972).
8 Blair-West, J. R., *et al.*, *Endocrinology*, **71**, 990 (1962).
9 Coghlan, J. P., *et al.*, *Proc. 4th Mtg. Int. Soc. Hypertension, Sydney, February 1976*, 22 (National Heart Foundation of Australia, 1976).
10 Sennett, J. A., *et al.*, *Circulation Res.*, **36**, Suppl. 1, 2 (1975).
11 Oliver, J. T., Birmingham, M. K., Butova, A., Li, M. P., and Chan, T. H., *Science*, **182**, 1249 (1973).

Mode of action of endogenous opiate peptides

AS THE culmination of a search for endogenous ligands of the opiate receptor[1-4], the structures of two closely related pentapeptides, isolated from porcine brain, with morphine-like properties in pharmacological tests, have been reported by Hughes *et al*[5]. The amino acid sequences of the pentapeptides, termed methionine enkephalin and leucine enkephalin, are Tyr-Gly-Gly-Phe-Met (Met^5-enkephalin) and Tyr-Gly-Gly-Phe-Leu (Leu^5-enkephalin), respectively. They behave as highly potent opiates in the mouse vas deferens and guinea pig ileum assays[5]. In addition, animals tolerant to morphine are also tolerant to enkephalin[6]. The action of the pentapeptides in the guinea pig ileum[5] and when injected intracerebrally or intracerebroventricularly[7-9] is very short lived, perhaps because of proteolytic degradation. A related peptide, with longer chain length has been shown to have a long duration of action[10]. Morphine and other narcotics have been shown to bind to the opiate receptor[11-13] and inhibit adenylate cyclase activity in homogenates of neuroblastoma × glioma hybrid cells[14,15], to inhibit cyclic AMP accumulation in intact hybrid cells[15,16] and in rat brain homogenates[17]. We have tested the effects of Met^5- and Leu^5-enkephalin on NG108-15 adenylate cyclase activity, and show here that endogenous opiate peptides are potent, receptor-mediated, inhibitors of adenylate cyclase of neuroblastoma × glioma hybrid cells.

NG108-15 hybrid cells were derived (unpublished results of B. Hamprecht, T. Amano and M.N.) by fusion of mouse neuroblastoma clone N18TG-2[18] with rat glioma clone

C6BU-1[19]. NG108-15 cells generate action potentials on electrical or chemical stimulation, synthesise acetylcholine (unpublished results of B. Hamprecht, T. Amano and M.N.), and form synapses with striated muscle cells[20]. They are richly endowed with opiate receptors[14].

The relationship between Met[5]-enkephalin, Leu[5]-enkephalin, and morphine concentrations and basal or prostaglandin E₁ (PGE₁)-stimulated NG108-15 adenylate cyclase activities are shown in Fig. 1a and b, respectively. The concentrations required for half-maximal inhibition of basal adenylate cyclase are 12, 40 and 1,500 nM, respectively, and 20, 120 and 1,500 nM for PGE₁-stimulated activity. Therefore, Met[5]-enkephalin and Leu[5]-enkephalin are approximately 100 and 25 times more potent than morphine as inhibitors of adenylate cyclase activity. The activities of Met[5]- and Leu[5]-enkephalin as inhibitors of adenylate cyclase are approximately equal to their activities as inhibitors of electrically evoked contractions of mouse vas deferens smooth muscle and are five times greater than their activities in a similar assay with guinea pig ileum[6].
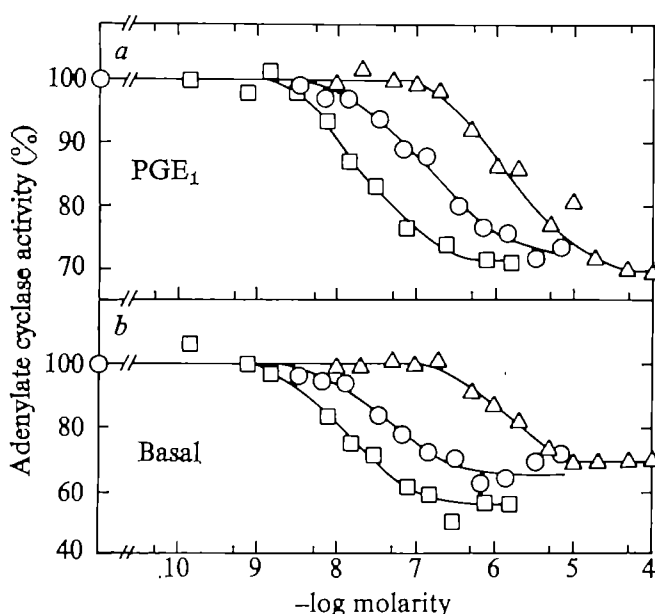
Fig. 1 Basal and PGE₁-stimulated adenylate cyclase activity of NG108-15 homogenates as a function of concentration of Met[5]-enkephalin (□), Leu[5]-enkephalin (○) or morphine (△). Reaction mixtures were incubated for 5 min at 37 °C with 97 μg of homogenate protein per tube. Other conditions are described in Table 1. Met[5]-enkephalin was synthesised by the rapid solid-phase procedure of Corley et al.[21] The peptide product (approximately 80% pure) was purified by high-pressure liquid chromatography using a reversed phase system (Corasil, Waters Associates) and an acetonitrile, 0.1-M acetic acid gradient Leu[5]-enkephalin, a synthetic product, was the gift of Drs R Simantov and S. H. Snyder.

The effectiveness of Met[5]-enkephalin (1.6 × 10⁻⁷ M) as an inhibitor of adenylate cyclase decreases during incubation (Fig. 2), and is lost after approximately 20 min. The degree of inhibition by morphine (2 × 10⁻⁷ M) also decreases during incubation, but to a lesser extent. Further work is needed to determine whether Met[5]-enkephalin and morphine are inactivated during incubation or whether an activator of adenylate cyclase is formed.

The effect of naloxone, an opiate antagonist which competitively inhibits narcotic binding to the opiate receptor, on Met[5]-enkephalin or morphine-dependent inhibitions of adenylate cyclase are shown in Fig. 3a and b, respectively. Different concentrations of naloxone were used in the presence of 10⁻⁸, 10⁻⁷ or 10⁻⁶ M Met[5]-enkephalin or in its absence. Naloxone completely reverses the inhibition of adenylate cyclase by each concentration of Met[5]-enkephalin tested. As expected for competitive interactions at a single

Table 1 Effect of Met[5]-enkephalin and morphine on adenylate cyclase activity in homogenates of neuroblastoma × glioma hybrid and parental cell lines

	Cell line		
	Hybrid	Neuroblastoma	Glioma
Additions	NG108-15	N18TG-2	C6BU-1
	pmol cyclic AMP per min per mg protein		
None	41.4	13.2	15.8
Met[5]-enkephalin	27.2	14.7	15.1
Morphine	35.9	16.8	16.8
PGE₁	217	115	20.5
PGE₁ + Met[5]-enkephalin	168	120	18.4
PGE₁ + morphine	164	121	20 2

Adenylate cyclase activity was measured by the procedure of Salomon et al.[13] with modifications described before[14]. Each reaction mixture contained, in a final volume of 100 μl: 30 mM Tris-HCl, pH 7.5, 5 mM MgCl₂; 160 mM sucrose; 20 mM creatine phosphate; 10 U of creatine kinase; 1 mM ³H-cyclic AMP (10,000 c.p.m.), 0.5 mM RO20-1724 (0.1% ethanol final concentration); 1 mM α-³²P-ATP (10⁶ c.p m.); and where indicated, 0.16 μM Met[5]-enkephalin; 20 μM morphine sulphate; and 10 μM PGE₁. In addition, 90, 95 and 118 μg of homogenate protein was present for NG108-15, N18TG-2 and C6BU-1, respectively. Tubes were incubated for 3 min at 37 °C. Reactions were stopped with 1 ml of 5% trichloroacetic acid containing 1 mM ATP and 1 mM cyclic AMP. ³²P-cyclic AMP and ¹⁴C-cyclic AMP (for recovery) were counted after purification as described by Salomon et al.[13]. Cell lines, growth conditions and the adenylate cyclase assay method have been described previously (refs 14 and 15 and unpublished results of B. Hamprecht, T. Amans and M. N.). Washed cells, suspended in 0.32 M sucrose, 0.01 M Tris-HCl at pH 7.5 were stored at −80 °C before use.

site, the amount of naloxone required to reverse the actions of Met[5]-enkephalin or morphine is a function of the affinity of the receptor for the ligands and their concentrations. The dissociation constant of naloxone, K_a, calculated[22] from the data in Fig. 3a and b is 3 × 10⁻⁸ M for reversal of Met[5]-enkephalin inhibition and 2 × 10⁻⁸ for reversal of morphine inhibition (see legend to Fig. 3). These values agree well with the dissociation constant of naloxone of 2 × 10⁻⁸ M previously determined by direct measurement of the binding of ³H-naloxone to the opiate receptor of the hybrid cells[14].

As Table 1 shows, Met[5]-enkephalin and morphine inhibit adenylate cyclase in homogenates prepared from NG108-15 hybrid cells, but not adenylate cyclase of the glioma parent, C6BU-1, which lacks narcotic receptors, or the neuroblastoma parent, N18TG2, which has fewer narcotic

Fig. 2 Time course of cyclic AMP formation by NG108-15 homogenates (97 μg per protein per tube) in the presence of 2 × 10⁻⁷ M morphine (○), 1.6 × 10⁻⁷ M Met[5]-enkephalin (●), or in the absence of added inhibitors (△). Assay conditions were as for Table 1.
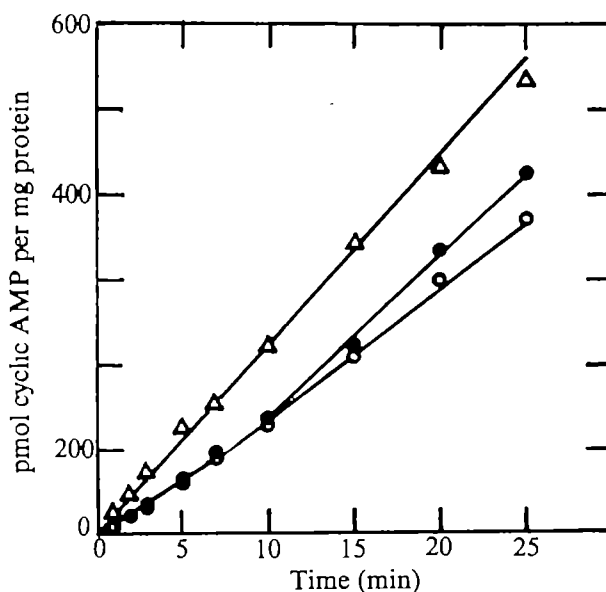
receptors than NG108–15 hybrid cells[14,15]. These data, together with the naloxone reversal of enkephalin inhibition, show that both opiate receptors and enkephalin are required for the inhibition of adenylate cyclase activity.

The apparent inhibition constants for adenylate cyclase of Met[5]- and Leu[5]-enkephalins as well as morphine and some peptides which are fragments of enkephalin are summarised in Table 2. The K_i for Met[5]-enkephalin is 12–20 nM. Leucine enkephalin is three to six times less active an inhibitor than Met[5]-enkephalin and the enkephalin fragments tested had little or no effect on adenylate cyclase activity. In other experiments, not shown here, the di-and tripeptides did not antagonise morphine-dependent inhibition of adenylate cyclase. Met[5]- and Leu[5]-enkephalins are somewhat more potent inhibitors of basal, than PGE₁-stimulated adenylate cyclase. Morphine usually behaves similarly, although not in the experiment shown here. These observations show that opiate and PGE₁ receptors are functionally coupled either by allosteric interactions or by competition for the same site on the enzyme.

Compound	K_i^* (nM)	
	Basal	PGE₁
Tyr–Gly–Gly–Phe–Met (Met[5]-enkephalin)	12	20
Tyr–Gly–Gly–Phe–Leu (Leu[5]-enkephalin)	40	120
Morphine	1,500	1,500
Tyr–Gly–Gly	>1,000,000†	—
Gly–Gly–Phe	>1,000,000‡	—
Phe–Leu	>1,000,000‡	—

Adenylate cyclase activity was determined as described in the legend of Table 1.
*Concentration of peptide required for half-maximal inhibition of basal or PGE₁ (10^{-4} M)-stimulated adenylate cyclase. Maximal inhibition of adenylate cyclase activity usually is 30–60% of total activity[14].
†Slight activity at 10^{-3} M (0–20% inhibition of adenylate cyclase).
‡No inhibition at 10^{-4} M.

An important question, that can be studied with cultures of NG108–15 cells, is whether the endogenous opiate peptides are addictive. As reported previously[28,29], NG108–15 cells and perhaps animals as well[7] chronically exposed to morphine, become tolerant to and dependent on morphine due to its dual action in inhibiting adenylate cyclase and eliciting a gradual increase in adenylate cyclase activity which compensates for the inhibition. Experiments are in progress to test whether hybrid cells become tolerant to and dependent on enkephalin and other opiate-like peptides.

Opiate peptides, 5–31 amino acid residues long[30–34], as well as β-melanocyte-stimulating hormone (β-MSH) apparently are derived by proteolysis of a common precursor, β-lipotropin[31], which had been isolated from pituitary[32]. Thus, β-lipotropin is a precursor of both an activator of adenylate cyclase (β-MSH)[33], and inhibitors of this enzyme (opiate peptides). Presumably, β-MSH and the endogenous opiates are recognised by different species of receptor which may be on the same or on different cells. Thus, positive and negative responses could result in separate cells or a single cell from different peptides derived from the same precursor.

The endogenous opiate peptides seem to be neurotransmitters or hormones which are destined for neurones with opiate receptors. The opiate peptide-receptor complex is a potent inhibitor of adenylate cyclase and thus the activations of adenylate cyclase by other species of neurotransmitters and hormones are suppressed. In effect, the opiate peptides act as pleiotropic desensitisers of many kinds of receptors which, in concert with the corresponding ligands, activate adenylate cyclase. Prolonged exposure to the opiate peptides may lead to an increase in adenylate cyclase activity as previously observed with morphine and other narcotics[28,29]. Thus, the opiate peptides may regulate the perception of incoming messages by neurones with opiate receptors in both a negative and positive manner.

We thank Christian B. Anfinsen, Lila Corley and Urs. Th Ruegg for help and advice with peptide synthesis and purification, Richard A. Streaty for assays and Doyle Mullinax for growing cells.

WERNER A. KLEE

Laboratory of General and Comparative Biochemistry,
National Institute of Mental Health,

MARSHALL NIRENBERG

Laboratory of Biochemical Genetics,
National Heart and Lung Institute,
National Institutes of Health,
Bethesda, Maryland 20014

1 Hughes, J., *Brain Res.*, **88**, 295–308 (1975).
2 Terenius, L., and Wahlstrom, A., *Acta physiol. scand.*, **94**, 74–81 (1975).
3 Pasternak, G. W., Goodman, R., and Snyder, S. H., *Life Sci.*, **16**, 1765–1769 (1975).

Fig. 3 *a*, relationship between naloxone concentration and the reversal of adenylate cyclase inhibition by 0.01 (△), 0.1 (□), or 1 μM (●) Met[5]-enkephalin. Also shown is the relationship between naloxone concentration and adenylate cyclase in the absence of Met[5]-enkephalin (○). Assays were for 5 min with 87 μg of homogenate protein per tube. *b*, a similar experiment without morphine (○), or with 1 μM morphine (●); or 100 μM morphine (▽). Reaction mixtures (90 μg of homogenate protein per tube) were incubated for 5 min. Other assay conditions are described in the legend of Table 1. Equilibrium binding constants (K_e) for naloxone were calculated by finding the concentration of naloxone at which the adenylate cyclase activity, in the presence of the highest concentration of inhibitor, is equal to that measured in the absence of naloxone at an inhibitor concentration 100 times lower. This concentration of naloxone is, by the dose-ratio method[15], equal to 99 times K_e.

Our experiments show that endogenous opiate peptides can be assayed rapidly by determining the opiate-peptide-dependent inhibition of adenylate cyclase which is reversed by naloxone. The range of the assay is 0.2–5.0 pmol of Met[5] enkephalin per 100 μl of reaction mixture or 0.1–2.5 pmol per 50 μl of reaction mixture. Thus, we have found, in collaboration with Goldstein and Cox, that a peptide which has been purified extensively from pituitary tissue[24] and has a higher molecular weight than enkephalin is a potent inhibitor of both basal and PGE₁-stimulated adenylate cyclase of NG108–15 cells. Enkephalin antagonists can be assayed in a similar manner.

4 Teschemacher, H., Opheim, K. E., Cox, B. M., and Goldstein, A., *Life Sci.*, **16**, 1771–1776 (1975).
5 Hughes, J., *et al.*, *Nature*, **258**, 577–579 (1975).
6 Waterfield, A. A., Hughes, J., and Kosterlitz, H. W., *Nature*, **260**, 624–625 (1976).
7 Belluzzi, J. D., *et al.*, *Nature*, **260**, 625–626 (1976).
8 Pert, A., Simantov, R., and Snyder, S. H., *Proc. natn. Acad. Sci. U.S.A.* (in the press).
9 Chang, J. K., Fong, B., Pert, A., and Pert, C. B., *Life Sci.* (in the press).
10 Feldberg, W. S., and Smyth, D. G., *J. Physiol., Lond.* (in the press).
11 Pert, C. B., and Snyder, S. H., *Science*, **179**, 1011–1014 (1973).
12 Terenius, L., *Acta Pharmac. Tox.*, **32**, 317–320 (1973).
13 Simon, E. J., Miller, J. M., and Edelman, I., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 1947–1949 (1973).
14 Klee, W. A., and Nirenberg, M., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 3474–3477 (1974).
15 Sharma, S. K., Nirenberg, M., and Klee, W. A., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 590–594 (1975).
16 Traber, J., Fischer, K., Larzin, S., and Hamprecht, B., *Nature*, **253**, 120–122 (1975).
17 Collier, H. O. J., and Roy, A. C., *Nature*, **248**, 24–27 (1974); *Prostaglandins*, **7**, 361–376 (1974).
18 Minna, J., Glazer, D., and Nirenberg, M., *Nature new Biol.*, **235**, 225–231 (1972).
19 Amano, T., Hamprecht, B., and Kemper, W., *Expl Cell Res.*, **85**, 399–408 (1974).
20 Nelson, P., Christian, C., and Nirenberg, M., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 123–127 (1976).
21 Corley, L., Sachs, D. H., and Anfinsen, C. B., *Biochem. biophys. Res. Commun.*, **47**, 1353–1359 (1972).
22 Kosterlitz, H. W., and Watt, A. J., *Br. J. Pharmac.*, **33**, 266–276 (1968).
23 Salomon, Y., Londos, C., and Rodbell, M., *Analyt. Biochem.*, **58**, 541–548 (1974).
24 Cox, B. M., Opheim, K. E., Teschemacher, H., and Goldstein, A., *Life Sci.*, **16**, 1777–1782 (1975).
25 Sharma, S. K., Klee, W. A., and Nirenberg, M., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 3092–3096 (1975).
26 Traber, J., Gullis, R., and Hamprecht, B., *Life Sci.*, **16**, 1863–1868 (1975).
27 Collier, H. O. J., Francis, D. L., McDonald-Gibson, W. J., Roy, A. C., and Saeed, S. A., *Life Sci.*, **17**, 85–90 (1975).
28 Guillemin, R., Ling, N., Burgus, R., *C.r. hebd. Séanc. Acad. Sci.*, Series D., **282**, 783–785 (1976).
29 Li, C. H., and Chung, D., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 1145–1148 (1976).
30 Bradbury, A. F., Smyth, D. G., and Snell, C. R., *Nature*, **260**, 793–795 (1976).
31 Bradbury, A. F., Smyth, D. G., and Snell, C. R., in *Peptides: Chemistry, Structure and Biology* (edit. by Walter, R., and Meienhofer, J.), 609–615 (Ann Arbor Science, Inc., 1975).
32 Li, C. H., *Nature*, **201**, 924 (1964).
33 Braun, T., Hechter, O., and Li, C. H., *The Endocrine Society*, Fifty Second Meeting, **48** (1970).

Molecular weights of antihaemophilic factor and von Willebrand factor proteins in human plasma

HUMAN factor VIII contains at least two biological activities, antihaemophilic factor (AHF) and von Willebrand factor (VWF). When highly purified, factor VIII is estimated to have a molecular weight of over 1.12×10^6, as determined by gel chromatography or sedimentation equilibrium centrifugation[1-4]. Some investigators[2-4] consider factor VIII a single glycoprotein with a number of covalently linked subunits. Since factor VIII can be dissociated in certain conditions, others[5-7] consider it to be a two-molecule complex consisting of a multi-subunit high molecular weight protein ($> 10^6$) with VWF activity that acts as a carrier molecule for a lower molecular weight (2.4×10^5) subunit with AHF activity. A third model suggests that both AHF and VWF are high molecular weight, separate proteins consisting of disulphide linked subunits[8,9].

These interpretations do not adequately explain certain clinical and laboratory observations. For example, von Willebrand patients infused with factor VIII preparations show a secondary rise in AHF activity when no measurable VWF activity or factor VIII-related antigen is present[10]. The molecular weight of the AHF proteins in this secondary rise, as determined by Agarose gel chromatography, showed that some AHF activity was associated with proteins with a molecular weight of at least 1×10^6 and some with proteins of considerably lower molecular weight; but neither protein showed VWF activity[11]. Furthermore, low molecular weight AHF has been found to aggregate with itself[8]. These findings could be interpreted more readily if native AHF and VWF proved to be two different molecules and the circulating AHF and VWF were low in molecular weight, comparable to that reported for the subunit size of reduced factor VIII.

Using rapid quantitative ultrafiltration of whole plasma to estimate the molecular weight of AHF and VWF and study their interactions, we determined that each native protein has a molecular weight of less than 300,000 and they exist in plasma as separate molecules.

Whole blood from healthy male and female adults was collected in 3.8% citrate (9:1). All subsequent processing steps were carried out at the specified experimental temperature.
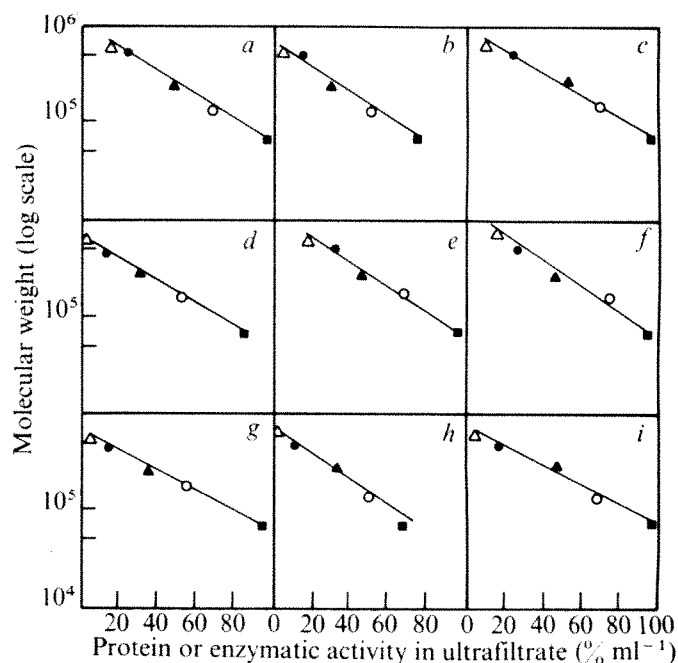
Fig. 1 Linear relationships between percentage marker proteins in filtrate and molecular weight on 9 different XM100 membranes labelled A–I. Each point represents the mean of at least 10 determinations. ■, Albumin; ○, gamma globulin; ▲, β-glucuronidase; ●, urease; △, β-galactosidase.

Plasma low in platelets was prepared by two centrifugations for 5 min at 16,000g and used immediately.

New Amicon 25-mm diameter XM100 ultrafiltration membranes were routinely cleaned by ultrasonication (Millipore Ultrasonic Cleaner Model #XX66 008 00) for 60 min in distilled water before use and by sonication for 15 min in distilled water before each application of proteins. Each filter was calibrated by ultrafiltration of several enzymes and other protein markers of known molecular weight and recalibrated before each subsequent use. The same filter could be used many times and between uses was stored at 4 °C in distilled water. Other methods for preparing and cleaning the membranes, including those suggested by the manufacturer, proved unsatisfactory and experiments in which they were used were, in general, not reproducible.

Solutions of each protein molecular weight marker containing 1.0 mg ml^{-1} (100%) were individually ultrafiltered and the

Table 1 Apparent molecular weight of native AHF and VWF determined by ultrafiltration at 37 °C on ultrasonicated Amicon XM100 membranes

Experiment	Molecular weight AHF	VWF
1	220,000	—
2	210,000	—
3	220,000	—
4	230,000	—
5	240,000	—
6	240,000	260,000
7	240,000	260,000
8	235,000	280,000
9	250,000	290,000
10	240,000	—
11	240,000	—
12	240,000	280,000
13	240,000	270,000
14	230,000	280,000
15	260,000	300,000
16	240,000	300,000
Mean ± s.d.:	236,000 ± 12,000	280,000 ± 15,000

The difference between the mean molecular weights calculated from the paired data is highly significant ($P < 0.001$).

filtrate was assayed for enzyme activity or protein concentration. The results were expressed as the percentage of the initial concentration per ml. The resulting calibration curves (Fig. 1) show a linear relationship between the percentage of protein recovered in the ultrafiltrate and the molecular weight. Each point represents the mean of 10 determinations on one of 9 different XM100 membranes. The precalibrated ultrasonicated membranes could thus be used to determine the molecular weights of various proteins from 68,000 to ~600,000, without loss of biological activity.

The first ml of ultrafiltrate was collected from 3.0 ml of fresh plasma filtered at 37 °C under nitrogen (10 pound/inch[s]). The filtrate and retentate were then examined for AHF activity (two-stage assay)[12], ristocetin VWF activity[13,14], and factor
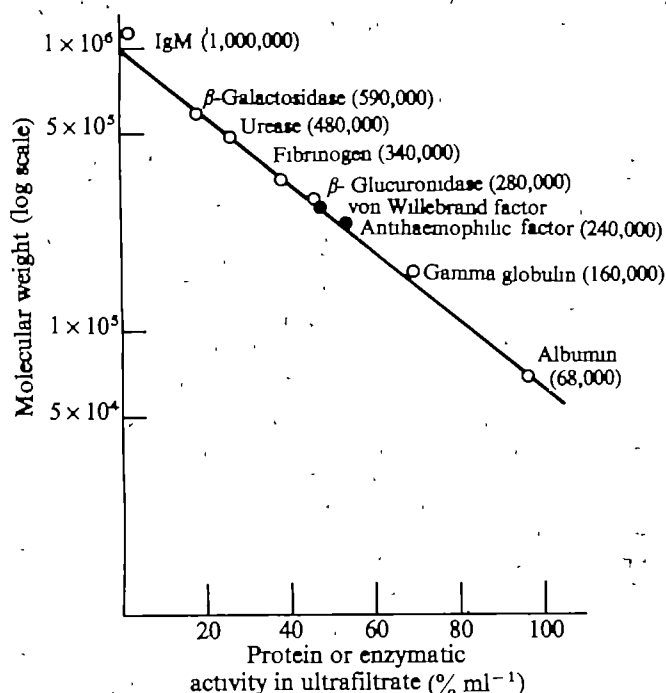
Fig. 2 Molecular weight of AHF and VWF determined from the percentage of fresh plasma in the ultrafiltrate in a single typical experiment using membrane E. Marker proteins and plasma were filtered through an XM100 membrane at 37 °C under nitrogen (10 pound/inch[r]).

VIII-related antigen (radioimmunoassay)[15]. Between the filtrate, the retentate and the material trapped on the filter support, virtually all the AHF or VWF activity in the sample could be accounted for.

In these conditions, 53% of AHF and 47% of VWF and antigen activity were recovered in the ultrafiltrate in a typical experiment. As evident from the calibration curve, the percentage of AHF or VWF in the ultrafiltrate corresponded to a molecular weight of 240,000 for AHF and 270,000 for VWF (Fig. 2). With use of a number of different calibrated membranes, a mean molecular weight of 236,000±12,000 for AHF and 280,000±15,000 for VWF was determined (Table 1).

The fact that the AHF and VWF in the 1 ml of filtrate was not a selected portion of the proteins in the 3 ml of sample applied was demonstrated when all 3 ml of plasma were filtered through a 43 mm instead of a 25 mm diameter membrane and similar molecular weight values were obtained.

To assess the effect of prolonged incubation on the filterability (molecular weight) of AHF and VWF, fresh plasma was incubated at 37 °C, and samples were assayed periodically. The proteolytic inhibitors Trasylol (50 kI Units ml⁻¹) and benzamidine (2×10⁻³ M) were added to prevent degradation of proteins during the experiment. After 6 h of incubation, the apparent molecular weight of AHF increased from 240,000 to 380,000 whereas the apparent molecular weight of VWF increased from 280,000 to more than 600,000. This increase was

attributed to aggregation of the AHF and VWF in vitro. The rate of aggregation proved to be faster at lower temperatures (25 or 4 °C).

Since the rate of aggregation appeared to increase with time and at low temperature, it was not surprising that the ultrafiltrates of cryoprecipitate and two commercial factor VIII preparations contained no measurable AHF or VWF activity, both of which were found in the retentate.

In summary, we have shown that the molecular weights of AHF and VWF in fresh human plasma are ~240,000 and 280,000. Their different molecular weights and rates of aggregation suggest that they are separate proteins, a conclusion consistent with other data from our laboratory using solid-phase polyelectrolyte fractionation. High molecular weight factor VIII seems to be an artefact arising from blood processing, low temperature storage, or fractionation procedures.

Quantitative ultrafiltration offers an important new method for the rapid, non-destructive estimation of the molecular weight of microgram quantities of proteins.

JACK NEWMAN
ROBERT B. HARRIS
ALAN J. JOHNSON

Department of Medicine,
New York University Medical Center,
550 First Avenue, New York,
New York 10016

[1] Bennett, B., Forman, W. B., and Rantoff, O. D., J. clin. Invest., 52, 2191–2197 (1973).
[2] Marchesi, S. L., Shulman, N. R., and Gralnick, H. R., J. clin. Invest., 51, 2151–2161 (1972).
[3] Legaz, M. E., Schmer, G., Counts, R. B., and Davie, E. W., J biol Chem., 218, 3946–3955 (1973).
[4] Shapiro, G A., Anderson, J. C., Pizzo, S. V., and McKee, A. P., J. clin. Invest., 52, 2198–2210 (1973).
[5] Weiss, H. J., Phillips, L. L., and Rosner, W., Thromb. Diath. Haemorrh., 27, 212–219 (1972).
[6] Owen, W. G., and Wagner, R. H., Thromb. Diath. Haemorrh., 27, 502–515 (1972).
[7] Rick, M. E., and Hoyer, L. W., Blood, 42, 737–747 (1973).
[8] Austen, D. E. G., Br. J. Haemat., 27, 89–99 (1974).
[9] Austen, D. E. G., Carey, M, and Howard, M. A, Nature, 253, 55–56 (1975).
[10] Bennett, B., Ratnoff, O. D, and Levin, J, J. clin. Invest., 51, 2597–2601 (1972).
[11] Bloom, A. L, Peake, I. R., and Giddings, J. C., Thromb. Res., 3, 394–404 (1973).
[12] Newman, J., Johnson, A. J., Karpatkin, M. H., and Puszkin, S., Br. J Haemat., 21, 1–20 (1971).
[13] Weiss, H. J., Hoyer, L. W., Rickles, F. R., Varma, A, and Rogers, J., J. clin. Invest., 52, 2708–2716 (1973).
[14] Macfarlane, D E., et al., Thromb. Diath. Haemorrh., 37, 306–308 (1975).
[15] Hoyer, L W., J. lab clin. Med, 80, 822–833 (1972)

N-terminal amino acid sequences of variant-specific surface antigens from *Trypanosoma brucei*

THE major pathogenic African (salivarian) trypanosomes can evade the immune response through a process of antigenic variation (reviewed in ref. 1). Mainly due to Gray's work[2,3], this variation is thought to be phenotypic—involving sequential expression of alternative genes and not resulting from contemporaneous mutation. The capacity for variation which exists within a clone, isolate or species, or within strains circulating within a geographical area, is uncertain (though probably extensive) and is relevant to an evaluation of the feasibility of developing vaccines. Although serotyping of trypanosome populations has been refined through the use of fluorescence methods[4,5], there is a need to characterise the relevant antigens and investigate the molecular basis of antigenic variation. Other workers[6–9] identified putative variable antigens and found them to be soluble glycoproteins[10–13], probably located in the surface coat[13,14]. The antigen preparations obtained, however, were generally heterogeneous, frequently unstable or in low yield. A recent report[15] described the complete purification in good yield of individual variant-specific surface antigens (VSSAs) from a series of clones of *Trypanosoma brucei*. The VSSAs seem to be the primary

Variants

	1	10	20	30

I   Thr-Asn-Asn-His-Gly-Leu-Lys-Leu-Gln-Lys-Ala-Glu-Ala-Ile-Cys-Lys-Met-Cys-Lys-Glu-

II   Ala-Lys-Glu-Ala-Leu-Glu-Tyr-Lys-Thr-Trp-Thr-Asn-His-Cys-Gly-Leu-Ala-Ala-Thr-Leu-Arg-Lys-Val-Ala-Thr-Gly-Val-Leu-Thr-Lys-Leu-Lys-Ser-His-Ile-

III   Thr-Asp-Lys-Gly-Ala-Ile-Lys-Phe-Glu-Thr-Trp-Glu-Pro-Leu-Gln-Leu-Leu-Thr-Gln-Asp-Phe-Gly-Asn-Leu-Tyr-Asn-Lys-Ala-Lys-$^{Lys}_{Asn}$-   -Asn-Leu-Asp-

IV   Ala-Glu-Ala-Lys-Ser-Asp-Thr-Ala-Ser-Gly-Ser-Val-Asn-Ser-Pro-Gln-Thr-Glu-Ala-Thr-Tyr-   -Ala-Gln-Leu-Ala-Lys-Thr-Leu-Gln-Ala-Leu-Asp-

Fig. 1 Amino terminal sequences of variant-specific surface antigens from four successive trypanosome variants. Sequences were obtained by automated liquid-phase and solid-phase techniques (see text). Phenylthiohydantoin derivatives were identified by high-pressure liquid chromatography on a column of microbondapak-C_{18} reverse-phase support (Waters Associates) using a Waters liquid chromatograph fitted with an ultraviolet detector. Solvent 1 was water–methanol–acetic acid (90:10:0.25), pH 4.1, and solvent 2 was water–methanol–acetic acid (10:90:0.025). A linear gradient of 95% solvent 1:5% solvent 2 to 40% solvent 1:60% solvent 2 was run at a flow rate of 2 ml min^{-1}. Normal sample quantity was 1–5 nmol, but where less sample was available, acetone was added to solvent 1 (50 µl l^{-1}) to give a flat baseline at high detector sensitivity. Identifications were confirmed by either thin-layer chromatography (all residues except His and Arg for positions 1–25), gas–liquid chromatography (non-polar residues) or amino-acid analysis after back-hydrolysis with HI (all residues except Trp and Met) (for details see ref. 19). Average repetitive yield was 93.5%. Residues in italics are interim assignments.

mediators of antigenic variation in *T. brucei*. Each antigenically homogeneous trypanosome clone yields one predominant glycoprotein antigen composed of a single polypeptide chain (apparent molecular weight 65,000 as determined by sodium dodecyl sulphate (SDS)–polyacrylamide gel electrophoresis) and, depending on the variant, 7–17% (w/w) carbohydrate (unpublished results of J. G. Johnson). The VSSA seems to be the major or sole constituent of the surface coat and comprises 5–10% of the total cell protein. Studies of the amino acid sequences, carbohydrate constitution and cell-surface organisation of several VSSAs are in progress, and we report here the N-terminal amino acid sequences of VSSAs isolated from four closely related variants.

Antigens were purified as described previously[15], from a series of clones prepared from cell populations isolated during 3 weeks from a single rabbit infected with a cloned strain (427) of *T. brucei*. N-terminal sequences (Fig. 1) were generally determined on approximately 8-mg amounts of intact S-^{14}C-carboxymethylated antigens using a dilute Quadrol programme[16] in a Beckmann 890C sequencer. Runs were performed in duplicate. The sequence of variant II was identical in two samples purified on separate occasions. Limited tryptic digestion (unpublished results of G. A. M. C. and J.G. Johnson) of variants I and III produced in each case a single large fragment of molecular weight 48,000 and 45,000, respectively (estimated from SDS–polyacrylamide gel electrophoresis). After purification on Ultrogel AcA 54 (variant III) or by isoelectric focusing (variant I), 1 mg of each tryptic fragment was attached to 3-aminopropyl glass using phenylene diisothiocyanate[17] and was degraded in an LKB 4020 solid-phase sequencer using a single cleavage programme (LKB Biochrom Ltd) similar to that described by Laursen[18]. The N-terminal sequences of these fragments from variants I and III were identical to the respective intact proteins through residues 2–15. The results for variants I and II extend and confirm those obtained previously using a micro solid-phase technique[17].

These results confirm the homogeneity of the purified antigens and demonstrate an absence of sequence homology in the N-terminal region. Sequential cleavage techniques generally permit identification of the initial 30–40 residues of an intact protein. This represents only 6% of the VSSA molecule. Future work, involving subfractionation of the antigen, will extend these sequences and may reveal homologies in other regions of the molecules. Other evidence (unpublished results of G. A. M. C. and J. G. Johnson) however, suggests that extensive variation occurs along most of the polypeptide chain, but that nevertheless there may be some conservation of tertiary structure, as indicated by the characteristics of selective tryptic cleavage of the native antigens. The isolation of large tryptic fragments with intact N-terminal regions containing several lysine residues from variants I and III suggests that the N-

terminus of the native VSSA is relatively protected from proteolytic attack. This observation, coupled with the reproducibility of sequences obtained with different preparations of VSSA, makes it likely that the reported sequences represent the true N termini of the intact antigens.

Except in respect of their lack of homology, the reported sequences are without intrinsically striking features. The amino acids are present in proportions which reflect the overall VSSA amino acid compositions. Even if the amino acid sequences of the VSSAs analysed in the study reported here are later shown to be atypically hypervariable, it seems that evolution of the genes responsible for successive variant types is a consequence of the accumulation of countless mutations. Our continuing structural studies are focusing on a search for regions of constant (or less variable) sequence which we hope will provide clues to the evolution of these antigens and also clarify the features necessary for their integration into the trypanosome surface membrane.

This work was performed in part at the Department of Microbiology, University of Alabama Medical Center, Birmingham, Alabama.

PAMELA J. BRIDGEN*
GEORGE A. M. CROSS

MRC Biochemical Parasitology Unit,
Molteno Institute,
University of Cambridge,
Downing Street,
Cambridge CB2 3EE, UK

JOHN BRIDGEN†

MRC Laboratory of Molecular Biology,
Hills Road,
Cambridge CB2 2QH, UK

[1] Vickerman, K., in *Parasites in the Immunized Host: Mechanisms of Survival, Ciba Foundation Symp.*, 25 new series, 53–70 (Associated Scientific Publishers, Amsterdam, 1974).
[2] Gray, A. R., *J. gen. Microbiol.*, 41, 195–214 (1965).
[3] Gray, A. R., *Trans. R. Soc. Trop. med. Hyg.*, 69, 131–138 (1975).
[4] Van Meirvenne, N., Janssens, P. G., and Magnus, E., *Ann. Soc. belge. Med. trop.*, 55, 1–23 (1975).
[5] Van Meirvenne, N., Janssens, P. G., Magnus, E., Lumsden, W. H. R., and Herbert, W. J., *Ann. Soc. belge. Med. trop.*, 55, 25–30 (1975).
[6] Williamson, J., and Brown, K. N., *Expl Parasit.*, 15, 44–68 (1964).
[7] Brown, K. N., and Williamson, J., *Expl Parasit.*, 15, 69–86 (1964).
[8] Le Page, R. W. F., Thesis, Univ. Cambridge (1968).
[9] Lanham, S. M., and Taylor, A. E. R., *J. gen. Microbiol.*, 72, 101–116 (1972).
[10] Allsopp, B. A., Njogu, A. R., and Humphryes, K. C., *Expl Parasit.*, 29, 271–284 (1971).
[11] Njogu, A. R., and Humphryes, K. C., *Expl Parasit.*, 31, 178–187 (1972).
[12] Allsopp, B. A., and Njogu, A. R., *Parasitology*, 69, 271–281 (1974).
[13] Vickerman, K., *J. Cell Sci.*, 5, 163–194 (1969).
[14] Vickerman, K., and Luckins, A. G., *Nature*, 224, 1125–1126 (1969).
[15] Cross, G. A. M., *Parasitology*, 71, 393–417 (1975).
[16] Brauer, A. W., Margolies, M. N. and Haber, E., *Biochemistry*, 14, 3029–3035 (1975).
[17] Bridgen, J. *Biochemistry*, 15, 3600–3605 (1976).
[18] Laursen, R. A., *Eur. J. Biochem.*, 20, 89–102 (1971).
[19] Bridgen, J., Graffeo, A., Karger, B. L., and Waterfield M., in *Instrumentation in Amino Acid Sequence Analysis* (edit. Perham, R. N.), 111–145 (Academic, London, 1975).

Enzymatic synthesis of deoxy-5-methyl-cytidylic acid replacing deoxycytidylic acid in *Xanthomonas oryzae* phage Xp12 DNA

PHAGE Xp12 on *Xanthomonas oryzae* was isolated from the irrigation water of a rice field. In the DNA of this phage, cytosine is completely replaced by 5-methylcytosine[1]. Isotope tracing studies on the biosynthesis of this unusual pyrimidine demonstrated that methylation of this cytosine is different from that of the 5-methylcytosine which is found in trace amounts in most plant and animal DNAs. The latter is methylated by methionine. On the other hand, the methyl group of 5-methylcytosine residues of Xp12 is derived from the 3-carbon of serine and not from the thiomethyl carbon of methionine[2]. In this investigation, the synthesis of deoxy-5-methylcytidylic acid (d5MCMP) was found to take place at the level of nucleotide, with deoxycytidylic acid (dCMP) and formaldehyde as substrates in the presence of 5,6,7,8-tetrahydrofolic acid (THFA) and an enzyme preparation from *X. oryzae* infected with Xp12 phage. This reaction seems to account for the synthesis of the unique pyrimidine, 5-methylcytosine, which occurs in the DNA of phage Xp12.

Exponential phase *X. oryzae* strain 604 grown in PS medium[3] (potato, 200 g; peptone, 5 g; sucrose, 15 g, Ca(NO₃)₂.4H₂O, 0.5 g; Na₂HPO₄.12H₂O, 2 g in 1 l of distilled water) was infected with a fivefold excess of Xp12, which has a latent period of 120 min in these conditions; aeration was continued for 40 min after infection. The cultures were rapidly chilled to 2 °C and the organisms were collected by centrifugation. After washing and resuspension in 0.05 M Tris-HCl buffer, *p*H 7.5, the cells were disrupted by using a Blacktone Model BP-2 Ultrasonic Probe. Unbroken cells and other particles were removed by centrifugation at 8,000*g* for 5 min. The supernatants were dialysed against 0.05 M Tris-HCl, *p*H 7.5 at 4 °C for 4 h. To show the induction by the phage of synthesis of the enzyme with methylating activity, protein synthesis was inhibited by the addition of chloramphenicol (30 μg ml⁻¹) just before phage infection. Extracts from uninfected and chloramphenicol-treated bacteria were prepared by the same method from parallel cultures. The protein concentration was approximately 13 mg ml⁻¹ as measured by the method of Lowry *et al.*[4].

The reaction mixture (150 μl) contained 50 μl enzyme extract, 10 μmol dCMP, 0.8 μmol THFA, 2 μmol ¹⁴C-formaldehyde (2 m Ci mmol⁻¹), 50 μmol MgCl₂.6H₂O, 32 μmol 2-mercaptoethanol and 40 μmol Tris-HCl, *p*H 7.5. The mixtures were incubated at 30 °C for 30 min. The reaction was stopped by adding 150 μl of 1.0 M sodium acetate, which dissociated 'active formaldehyde' to formaldehyde and tetrahydrofolate[5], followed by 100 μl of 0.4 M dimedone (in 50% ethanol) to precipitate formaldehyde[6]. The mixture was heated in boiling water for 5 min and the precipitated protein and formaldehyde were removed by centrifugation at 8,000*g* for 5 min. The supernatant and authentic d5MCMP, prepared from Xp12, were applied respectively to Whatman No. 1 chromatography paper and developed with isopropanol–HCl–H₂O (65: 16.6: 18.4).

Table 1 shows that the extract of infected cells promotes the synthesis of d5MCMP. A sharp peak of radioactivity was detected at the position corresponding to authentic d5MCMP. For general enzyme assays, the area round the peak was cut out from the paper, placed in a vial, and the radioactivity counted with a Packard Model 3375 scintillation spectrometer. To identify the reaction product, the spot containing the radioactivity was cut out and eluted with water. The eluate was dried and hydrolysed with 50 μl of 90% formic acid at 175 °C for 30 min. Acid hydrolysis of the nucleotide yielded a base with *R*f identical to that of authentic 5-methylcytosine in three solvent systems: isopropanol–HCl–H₂O (170 : 40 : 39), *n*-butanol–H₂O (86 : 14) and *n*-butanol–H₂O (86 : 14) with 5% NH₄OH in air. For further identification, a small amount of the substance obtained by evaporation of the eluate from a chromatogram spot was deaminated with HNO₂. Deamination was carried out by adding 50 μl of 2 M NaNO₂ and 10 μl of glacial acetic acid to the eluate. After standing overnight at 25 °C, the solution was applied directly to Whatman No. 1 paper. A sample of authentic 5-methylcytosine was treated in the same way. Both the unknown and authentic 5-methylcytosine were found to convert into a substance identical to thymine in its chromatographic properties.

On the basis of these observations, we conclude that an enzyme found in the infected bacteria catalyses the direct methylation of dCMP, but that this enzyme, deoxycytidylate methyltransferase, is absent from uninfected cells. The reaction required the presence of THFA and Mg²⁺, in addition to dCMP and formaldehyde as substrates. Deoxy-5-methylcytidylic acid can thus be synthesised in phage Xp12-infected *X oryzae* by the following reaction

$$\text{dCMP} + \text{HCHO} \xrightarrow{\text{THFA, Mg}^{2+}} \text{d5MCMP}$$

TSONG-TEH KUO
JENN TU

Institute of Botany,
Academia Sinica,
Taipei, Republic of China

1 Kuo, T. T., Huang, T. C., and Teng, M. H., *J. Molec. Biol.*, **34**, 373–375 (1968).
2 Ehrlich, M., Ehrlich, K., and Mayo, J. A., *Biochim. biophys. Acta.*, **395**, 109–119 (1975).
3 Kuo, T. T., Huang, T. C., Wu, R. Y., and Yang, C. M., *Bot. Bull. Acad. sin., Taiwan*, **8**, 246–254 (1967).
4 Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J., *J. biol. Chem.*, **193**, 265–275 (1951).
5 Osborn, M. J., Talbert, P. T., and Huennekens, F. M., *J. Am. Chem. Soc.*, **82**, 4921–4927 (1960).
6 Horning, E. C., and Horning, M. G., *J. org. Chem.*, **11**, 95–99 (1946).

Table 1 Incorporation of H¹⁴CHO into deoxy-5-methylcytidylic acid by extracts of Xp12 infected *X. oryzae*

Reaction mixture	Specific activity (nmol d5MCMP formed per mg protein in 30 min)
Complete system*	20.0
Complete system − Mg²⁺	15.7
Complete system − THFA	1.5
Complete system − dCMP	0.6
Reagent mixture + boiled enzyme†	0.8
Reagent mixture + extract of chloramphenicol-treated cells‡	0.8
Reagent mixture + extract of uninfected cells	1.6

*Complete system as described in text.
†Extract from infected bacteria was heated for 5 min in boiling water.
‡Extract from infected bacteria treated with chloramphenicol. Chloramphenicol was added just before phage infection.

Definitions of free energy levels in biochemical reactions

IN considering the energetics of an enzyme-catalysed reaction it is of interest to enquire into what constitutes the "drive", in a thermodynamic sense, for the reaction, particularly when the enzyme is involved in an energy-coupled process. This often

arises when speculations are made on the operation of a system on the basis of equilibrium constants for the reaction scheme determined on the system, or fragments of it, *in vitro*. In the case of active transport or muscular contraction the free energy for transport or contraction is derived from the hydrolysis of ATP and the overall drive arises because the chemical potential of ATP is greater than that of its products (ADP and P_i) at physiological concentrations. But to understand how the drive is expressed in the mechanism of such processes it is necessary to consider and define the relative free energy levels of the intermediates of the reaction. The free energy and standard free energy as normally defined as not sufficient for this purpose. They both, in general, refer to the system as a whole and are used in the context of free energy changes involving substrates, products and ligands. The drive of the system, in the sense used here, is to be sought in the relative disposition of the free energy levels of an individual (time-averaged) macromolecule and in the manner in which these levels favour (stochastically) a particular direction of the flux in what may be a complicated reaction scheme. We refer to these free energy levels of individual macromolecules as the "basic" free energy levels. To avoid confusion we have termed the overall free energy change of the whole system for a given transition as the "gross" free energy change. The definitions of and relationships between the standard, basic, and gross free energy levels for a system of macromolecules (either in free solution or immobilised) are given below, together with a simple example relevant to muscular contraction. A more mathematically detailed version of this work with more complicated examples for both muscle and active transport has been published elsewhere[1,2].

We consider a macromolecule (for example, enzyme) which can exist in a number of states (for example, enzyme-substrate, enzyme-product). In free solution, dilute with respect to enzyme, the chemical potential of the *i*th state can be written (per molecule and neglecting the activity coefficient) as

$$\mu_i = \mu^0_i + kT \ln c_i \qquad (1)$$

In the case of active transport and of muscle, the macromolecule is bound and the chemical potential can be derived from statistical mechanics[1] as

$$\mu_i = A_i + kT \ln p_i \qquad (2)$$

where A_i is the Helmholtz free energy of one macromolecule in the *i*th state and p_i is the probability of occupancy of this state ($\sum_i p_i = 1$). As this expression is the more general we shall use it throughout, noting for the free solution case that $c_i = bp_i$ and $A_i = \mu^0_i + kT \ln b$, where b is a constant. (In the bound macromolecule case, A_i is essentially the same as the Gibbs free energy G_i since the two quantities differ only by a negligible pV_i term.)

The free energy change for a transition between two states *i* and *j* of a macromolecule E can now be defined (for a kinetic scheme which may have a large number of steps and hence states of *E*). For generality, consider a transition involving the binding of a ligand (or substrate, and so on)

$$\begin{array}{c} \alpha^*_{ij} \\ E + L \rightleftharpoons EL \\ \alpha_{ji} \end{array} \qquad (3)$$

where E and EL are the *i*th and the *j*th states respectively, α^*_{ij} is

the (second order) forward rate constant and α_{ji} is the (first order) reverse rate constant. (The asterisk replaces the prime used previously[1-4] to denote second order quantities, to prevent confusion with the different assignment of the latter as below.)

The standard free energy change is

$$\Delta\mu^0_{ij} = A_i + \mu^*_L - A_j \qquad (4)$$

Note that we use Δ in the unconventional sense (initial) − (final), to make most such quantities used below positive.

The basic free energy change is

$$\Delta A'_{ij} = A_i + \mu_L - A_j \qquad (5)$$

The gross free energy change is

$$\begin{aligned} \Delta\mu'_{ij} &= \mu_i + \mu_L - \mu_j \\ &= \Delta A'_{ij} + kT \ln(p_i/p_j) \end{aligned} \qquad (6)$$

The meaning given to the prime here is as follows: A_i and A_j can be considered as the free energy levels of individual macromolecules, but where binding of a ligand occurs one of them must be "corrected" for the actual (not "standard") ligand concentration if they are to be compared[3,4]. Thus we can write $A'_E = A_E + \mu_L$, or more generally $A'_i = A_i + \mu_L$, and the use of the prime implies that the As have been "corrected" where appropriate.

The standard free energy change can be determined if the conventional equilibrium constant ($K^*_{ij} = \alpha^*_{ij}/\alpha_{ji}$) for the transition is known, since

$$\Delta\mu^0_{ij} = kT \ln K^*_{ij} \qquad (7)$$

Note that the asterisks refer to the "second-order" quantities.

It is worth noting that for most biological reactions in free solution the standard and gross free energy changes are, of course, simply related to the corresponding standard and "actual" molar Gibbs free energy changes, that is, $-\Delta G^0_{ij} = N_0\Delta\mu^0_{ij}$ and $-\Delta G_{ij} = N_0\Delta\mu'_{ij}$, where $N_0 \equiv$ Avogadro's number and the Δ on the left in both cases is the customary $\Delta \equiv$ final − initial. We approach these changes, however, from the unconventional point of view of single transitions between free energy levels of macromolecular states on a pseudo-isomeric basis and it is appropriate to consider the changes on a "per molecule" basis.

It is easily shown that the basic free energy change is given by

$$\Delta A'_{ij} = kT \ln c_L K^*_{ij} = kT \ln K_{ij} = kT \ln(\alpha_{ij}/\alpha_{ji}) \qquad (8)$$

where $K_{ij} = c_L K^*_{ij}$, and $\alpha_{ij}(= c_L\alpha^*_{ij})$ is the pseudo-first-order rate constant for the transition in question.

Thus, "correcting" for ligand concentration yields first-order rate constants and the corresponding (basic) equilibrium constants, which can be treated as if the transition occurred between isomers of the macromolecule.

Since the basic free energy change refers to the difference between the free energies of states of a single macromolecule,

and hence influences its stochastic behaviour, it is of fundamental importance in determining the drive in a reaction sequence. A diagram of the relative basic free energy levels (see Fig. 1) is, as shown above, closely related to the kinetic description in terms of first-order rate constants (though energy barriers between states are not shown in Fig. 1). Thus a downwards change favours flux through the system in the direction of the

basic change by a term $(kT \ln p_i/p_j)$ in equation (6). This is a typical "concentration" term of purely entropic origin. A situation may thus arise where the basic free energy change between two successive states is small, but the gross free energy change is large. (This can occur in an enzymatic cycle when the transition $i \rightarrow j$ is the rate-limiting step, so that in the steady state $p_i \gg p_j$. An example of this is given below.)
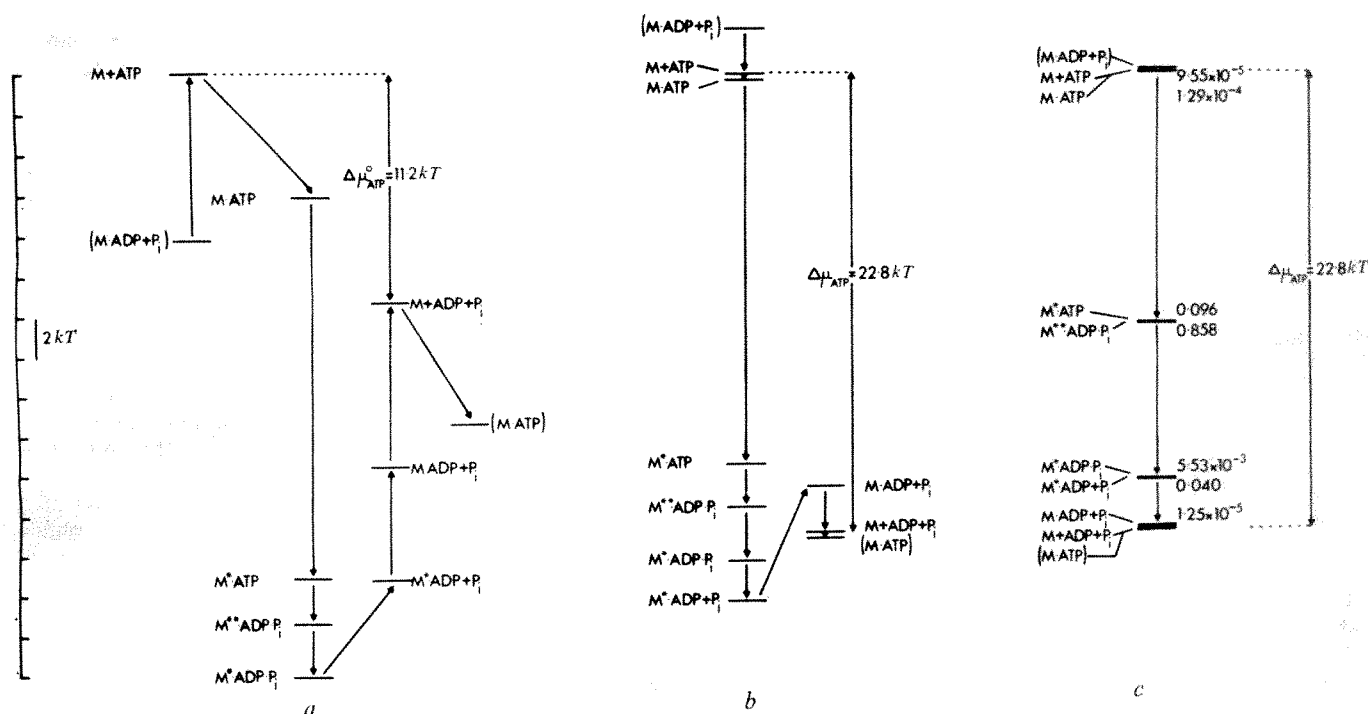
Fig. 1 Free energy levels for the intermediates in the hydrolysis of ATP by myosin *S1* (*M*) at 20 °C. Concentrations of ATP, ADP and P_i were chosen as 3 mM, 0.03 mM and 1 mM respectively. $\Delta\mu^0{}_{ATP}$ was taken as 11.2 kT. M + ATP was arbitrarily set at the same free energy level in (*a*), (*b*) and (*c*). Intermediates in parentheses belong to sets of levels immediately above or below those shown in the figure. The numbers opposite gross free energy levels (*c*) are the steady-state occupancies (p_i).

overall drive (ATP→ADP+P$_i$) whereas an upwards change opposes it. It is the downward basic changes that constitute the stochastic drive at the level of the intermediates in a reaction.

As for the standard free energy, in the case of transitions where the binding of ligands, and so on is not involved (isomeric transitions such as $ES \rightleftharpoons EP$), the standard free energy change is easily seen to be numerically identical to the basic change. But in non-isomeric transitions the two are by no means the same; in fact, the standard free energy change is of little interest as it refers to ligand, and so on at molar (that is "standard") concentration. Similarly the basic (isomeric) equilibrium constant K_{ij} is more relevant to actual conditions than the "standard" equilibrium constant $K^*{}_{ij}$.

The gross free energy change can be calculated when the probabilities (or concentrations) of the intermediates are known, in addition to the equilibrium constants. It gives the Gibbs free energy change that actually occurs in a transition when the whole system is considered. This is to be contrasted with the basic change description in which a single macromolecule undergoes transitions stochastically[5] and is "unaware" of the rest of the system. Thus the gross change is related to the overall flux and entropy production in a transition[1,2]. It should be noted that, for given substrate, and so on, concentrations, the basic changes and levels are invariant whereas the gross changes depend on the exact conditions of the system (in the steady state or during transients).

The gross free energy change for a transition differs from the

Suppose now that these states are coupled to another reaction, to transport or, in the case of muscle, to the performance of mechanical work; it might be deduced that the coupled process is being "driven by the concentration term". This would, in general, be incorrect; the efficiency and flux of a coupled process that is far from equilibrium result from the detailed kinetics of the entire cycle of the driving reaction and the kinetic constraints imposed on it by the coupled process.

For an enzyme which hydrolyses ATP, in each completed cycle the decrease in free energy is given by the sum of either the basic or the gross changes, that is

$$\Delta\mu_{ATP} = \Sigma\Delta A'_{ij} = \Sigma\Delta\mu'_{ij} > 0 \qquad (9)$$

The free energy decrease per molecule of ATP hydrolysed is given by

$$\Delta\mu_{ATP} = \Delta\mu^0{}_{ATP} + kT \ln(c_{ATP}/c_{ADP}c_{P_i}) \qquad (10)$$

Here again the standard free energy change is of no direct interest.

A comparatively simple example related to muscular contraction will now be given to illustrate the different definitions. (More complicated models both of active transport and of muscle are discussed in refs 1–3). The kinetic scheme for the hydrolysis of ATP by myosin subfragment 1 (M) has been determined[6–8] as

$$M + ATP \xrightleftharpoons[2.2 \times 10^5 s^{-1}]{10^8 M^{-1} s^{-1}} M \cdot ATP \xrightleftharpoons[2.3 \times 10^{-6} s^{-1}]{400 s^{-1}} M^* \cdot ATP \xrightleftharpoons[17.8 s^{-1}]{160 s^{-1}} M^{**} \cdot ADP \cdot P_i$$

$$3.8 \times 10^{-3} s^{-1} \Big\Updownarrow 0.06 s^{-1}$$

$$M + ADP \xrightleftharpoons[2.7 \times 10^4 s^{-1}]{10^8 M^{-1} s^{-1}} M \cdot ADP \xrightleftharpoons[1.4 s^{-1}]{400 s^{-1}} M^* \cdot ADP \xrightleftharpoons[7.3 \times 10^5 s^{-1}]{10^8 M^{-1} s^{-1}} M^* \cdot ADP \cdot P_i$$
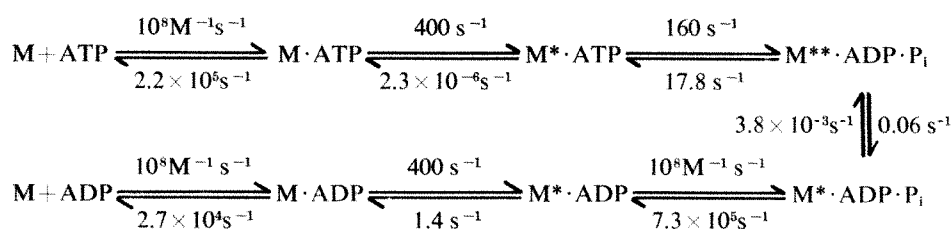
where the rate constants for binding ATP, ADP and P_i have been chosen as $10^8 \, M^{-1} s^{-1}$ and the forward rate constant for the third step represents the minimum likely value. Assuming reasonable values for the physiological concentrations of ATP, ADP and P_i, the relative free energy levels between states can be calculated, taking the steady-state case for the gross free energy changes (Fig. 1).

As noted above, the magnitude of free energy changes for the standard and basic diagrams coincide only for isomeric transitions (for example, $M \cdot ATP \rightleftharpoons M^* \cdot ATP$). As regards the difference between the basic and gross diagrams, the former has both upward and downward transitions, the principal downward changes being for the two-step binding of ATP. The free energy changes for the gross diagram are all downward. This is a natural consequence of the steady state in a single cycle in which the flux through the system is unidirectional[1]. The largest free energy change is for the two-step binding of ATP in both the basic and gross diagrams, but in the gross diagram there is not inconsiderable change at the rate-limiting step ($M^{**} \cdot ADP \cdot P_i \rightleftharpoons M^* \cdot ADP \cdot P_i$). This is a consequence of the build-up of the occupancy of the state preceding the rate-limiting step. In the gross diagram several pairs of states are at nearly the same free energy level indicating that, in the steady-state conditions used, these pairs are essentially in equilibrium (indeed, when this is considered in conjunction with the sparsity of occupation of the states M, $M \cdot ATP$ and $M \cdot ADP$, the scheme reduces essentially to a two-state energetic and kinetic process consisting of $M^* \cdot ATP$ "lumped" with $M^{**} \cdot ADP \cdot P_i$ and $M^* \cdot ADP \cdot P_i$ "lumped" with $M^* \cdot ADP$). In all the diagrams, for each cycle completed, the subsystem or system progresses to another, identical, set of levels shifted downwards from the previous set by $\Delta\mu_{ATP}$ ($\Delta\mu^0_{ATP}$ for the standard diagram).

The S1 scheme provides considerable scope for speculation about the basic free energy levels of the cyclic interaction of myosin with actin in the crossbridge mechanism of muscular contraction. Unfortunately full details of the corresponding scheme in the presence of actin are not yet known; also there are serious difficulties in translating a kinetic scheme derived for free solution to that for the intact myofibrillar matrix[10]. Nevertheless, modifying slightly a much simplified scheme for actomyosin[11],

$$M^{**} \cdot ADP \cdot P_i \rightleftharpoons M^* \cdot ATP$$

$$\Updownarrow \, \circlearrowright \, \Updownarrow$$

$$AM \cdot ADP \cdot P_i \rightleftharpoons AM$$
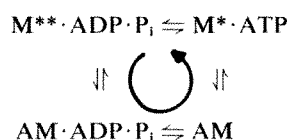
the M and AM states correspond respectively to detached and attached states of crossbridges in muscle and the drive (and directionality) of contraction must be reflected in the downward progression of basic free energy levels between successive AM states in the (biochemical) actomyosin scheme[3,12]. Large downward changes elsewhere in the cycle could lead to a waste of free energy in the gross free energy representation for intact muscle. It seems that such changes may not exist, at least in the basic representation for the actomyosin scheme (H. D. White and E. W. Taylor, in preparation); for example, the dissociation constant and hence the basic free energy change for $AM + ATP \rightleftharpoons A + M^* \cdot ATP$ is small, the large binding energy for ATP to myosin being offset by that of actin to

myosin[13]. Furthermore, the basic change for $M^* \cdot ATP \rightleftharpoons M^* \cdot ADP \cdot P_i$ is comparatively small (Fig. 1). It is therefore possible to regard $M^{**} \cdot ADP \cdot P_i$ as a "high energy state", in relation to the last attached state (AM) in the scheme above, in the sense that the (probably) large downward basic free energy change between them serves as the drive for contraction (although it seems unlikely that all the available energy from ATP hydrolysis is conserved in the state $M^{**} \cdot ADP \cdot P_i$; compare ref. 14).

Gray[15] has argued on quantum mechanical grounds for an irreversible step in the crossbridge cycle. This invites the comment that it all depends on what you mean by irreversible; as we have argued above, it does seem that a necessary requirement for a high efficiency in muscle is a large downward basic free energy change from the state prior to attachment to the last attached state (measured in the latter in the absence of mechanical constraints), but that does not necessarily mean that the transitions would be entirely irreversible. However, a full treatment of this kind of problem involves the averaging of rate constants and free energy levels over the distribution of crossbridges in each attached state[3]. The actual system considered (for comparison with experiment) is not a single crossbridge but a macroscopic ensemble of crossbridges under a potentially wide distribution of mechanical constraints; there is no simple way of introducing these constraints, for example, by introducing a work-dependent opposing thermodynamic flux into the scheme (compare ref. 16). We hope to return to this problem in a future paper.

Finally, it should be stressed that a basic free energy diagram gives no more than a useful summary of the energetic situation for a single macromolecule. This may be useful in model-building[3] or in comparing different systems, but a full description of the energetics of a system in particular conditions requires a detailed knowledge of the kinetics of the system to generate (and understand) the gross free energy diagram.

This subject is treated in more detail and in a much more general context elsewhere[9].

R. M. SIMMONS
TERRELL L. HILL

*Department of Physiology,
University College London,
Gower Street, London WC1E 6BT, UK
and Laboratory of Molecular Biology,
National Institute of Arthritis,
Metabolism and Digestive Diseases,
National Institutes of Health,
Bethesda, Maryland 20014*

1 Hill, T. L., and Simmons, R. M., *Proc. natn. Acad. Sci. U.S.A.*, 73, 95 (1976).
2 Hill, T. L., and Simmons, R. M., *Proc. natn. Acad. Sci. U.S.A.*, 73, 336, 2165 (1976).
3 Hill, T. L., *Prog. Biophys. molec. Biol.*, 28, 267 (1974).
4 Hill, T. L., *Biochemistry*, 14, 217 (1975).
5 Hill, T. L., and Chen, Y., *Proc. natn. Acad. Sci. U.S.A.*, 72, 1291 (1975).
6 Bagshaw, C. R., Eccleston, J. F., Eckstein, F., Goody, R. S., Gutfreund, H., and Trentham, D. R., *Biochem. J.*, 141, 351 (1974).
7 Mannherz, H. G., Schenck, H., and Goody, R. S., *Eur. J. Biochem.*, 48, 287 (1974).
8 Trentham, D. R., Eccleston, J. F., and Bagshaw, C. R., *Quart. Rev. Biophys.*, 9, 217 (1976).
9 Hill, T. L., *Free Energy Transduction in Biology* (Academic Press, New York, in the press).
10 Hill, T. L., and Eisenberg, E., *Biochemistry*, 15, 1629 (1976).
11 Lymn, R. W., and Taylor, E. W., *Biochemistry*, 10, 4617 (1971).
12 Huxley, A. F., and Simmons, R. M., *Nature*, 233, 533 (1971).
13 Marston, S., and Weber, A., *Biochemistry*, 14, 3868 (1975).
14 McClare, C., *Nature*, 240, 88 (1972).
15 Gray, B. G., *Nature*, 253, 436 (1975).
16 Taylor, E. W., *Current Topics in Bioenergetics*, 5, 201 (1973).

reviews

Introducing HLA

Heather M. Dick

The HLA System: An Introductory Survey. (Monographs in Human Genetics, Vol. 7.) By A. Svejgaard *et al.* Pp. viii + 103. (Karger: Basel, London and New York, 1975.) SF/DM 38; $14.75.

THIS is an excellent little volume which sets out with a refreshing lack of pretentiousness to give a simple introduction to the HLA system. The authors certainly achieve their main purpose and offer a text which will be invaluable to those who wish to gain access to exciting new developments.

The HLA antigens are no longer the specialised subject of a few devoted transplantation workers. The wide range of problems in biology and clinical medicine which can be related to the major histocompatibility system is evidence of the explosion of knowledge which has occurred in this subject.

After a clear exposition of the components of the HLA system, in particular the genetics, we are led straight into the most recent and exciting work on the similarities between human and mouse immune response mechanisms. The section on biochemistry is brief— reflecting the paucity of information on this aspect until very recently, but introducing the reader immediately to newer and more rewarding developments on the structure of HLA substance. The practical applications of studies on HLA antigen frequencies are dealt with adequately enough for the general reader.

The authors are perhaps too modest in not including more of their own excellent work on the statistical and technical problems of estimating associations between specific HLA antigens and diseases. They also shy away from a detailed discussion of the mechanisms which might be responsible for such asociations. Although they list several impeccable reference sources dealing with this problem, it would probably be helpful to readers who are not intimately involved in the current work to have the arguments for and against the suggested hypotheses presented in some detail. Perhaps the authors felt they had already made it obvious that their preference was for the genetically de-

termined "immune-response gene" hypothesis, but it would not weaken their argument to expose the other hypotheses of molecular mimicry, or HLA as virus receptor to the cold light of print.

The final section on methodology is adequate to provide background information for those not directly concerned with technique, although the difficulties of explaining MLC typng to an uninitiated audience are painfully obvious. There is a very full bibliography to complete the text.

The materal presented is up-to-date and the authors are to be congratulated

on presenting a useful and readable volume which includes so much contemporary material. Textual misprints are a minor source of annoyance. The occasional lapses from colloquial English by the team of Danish authors might have been edited out, but do not detract in any way from the value of the book. Certainly, this reviewer could not hope to do as well in Danish. □

Heather Dick is a Consultant in Clinical Immunology at Glasgow Royal Infirmary, in charge of histocompatability services for Glasgow and the west of Scotland.

Hückel model

The HMO Model and Its Application. Vol. 1: Basis and Manipulation. By Edgar Heilbronner and Hans Bock. Pp. xv + 454. (Wiley: London and New York; Verlag Chemie: Weinheim, February 1976.) £16; $35.20.

THIS is the first volume of a three-volume treatise which is intended to acquaint "the practically oriented chemist having a relatively modest mathematical background" with the principles of the Hückel molecular orbital (HMO) model and with the simple techniques with which it can be used to obtain approximate but very useful qualitative results without recourse to a computer. Volume 1 contains all the theory and applications, and 250 problems; volume 2 will contain the solutions to the problems; and volume 3 will contain HMOs and derived quantities for simple π-elecfron systems.

This all seems splendid, but I cannot recommend the practically oriented chemist to rush out for his copy, and I doubt whether he would if I did. One reason is that this volume alone costs £16, so that the whole set is going to add up to a lot of money. Furthermore, for all that the authors claim to "fit the intuitive conceptual world of the chemist", there is not really very much discussion of the chemistry. Coulomb integrals, for instance, appear as numbers, with due acknowledgement to Streitwieser but no attempt to relate

their values to the chemist's notion of electronegativity, which indeed is not mentioned anywhere. More seriously, the treatment of reactions is now badly out of date, the book being an unrevised translation of the 1968 German edition; it is a great pity that the opportunity was not taken to add, for example, an account of the frontier orbital method, which has proved so fruitful in recent years.

Other cricitisms, which I make with some reluctance because I am in sympathy with the authors' intentions and find a great deal of useful material here, are that the style is intensely turgid and polysyllabic, and that some mathematical derivations are given in painful detail, whereas other results are quoted without the proof which might quite easily have been given. Finally, I disagree myself with the authors' policy of saying very little indeed about the deficiencies of the HMO model and about the way they can be overcome in more elaborate methods.

I do not underestimate the difficulty of putting molecular orbital theory over to the non-mathematical chemist. This is not a bad book by any means, but it is too heavy, too expensive and too late.

A. J. Stone

Dr Stone is an Assistant Director of Research at the Chemical Laboratories, University of Cambridge, UK.

Cell handbook

Cell Biology. (Biological Handbooks, 1.) Compiled and Edited by Philip L. Altman and Dorothy Dittmer. Pp. xix+454. (Federation of American Societies for Experimental Biology: Bethesda, Maryland, 1976.)

FROM time to time a new book appears which is not simply a rehash of those that have gone before. Until now, biology has lacked its "Beilstein" and it is often a time-consuming business to find some small piece of information which you know has been published. This book compresses the maximum amount of factual information into the minimal space and at the same time gives key references from which a quick general overview of the field of interest may be made To give some examples, do you want to know what nutritional mutants of mammalian cells are available; 31 types are listed and 40 references are given to the field, all in the space of 2 pages. Are you interested in cloning mammalian cells; 52 examples are given and 49 references in only 3 pages. How long would it take to find the oxygen consumption of 618 different mammalian tissues; it is all here

with 198 references compressed into 17 pages. As this one volume extends to 400 pages, excluding appendices and a full index, the amount of information available is remarkable.

Some sample headings, similar in range to those listed above, will give an indication of the scope of this miniature encyclopaedia: resting cell membrane protentials, rheological properties of cells, changes in cell surface membranes with malignant transformation, permeability of inner mitochondrial membrane to anionic substrates, gross chemical composition of microsomes, occurrence of microfilaments in non-muscle cells and tissues, DNA content of cell nuclei. Obviously this is not a book that can be read but it is a book which most biologists are likely to treasure. Inevitably the choice of key references is open to criticism and at times the choice seems somewhat parochial. Nevertheless sufficient are given to enable the research worker to get off to a flying start

This series is likely to be a must in many libraries. **T. S. Work**

Dr Work is Head of the Division of Biochemistry at the National Institute for Medical Research, London, UK.

Somatic cell genetics

Variation, Senescence and Neoplasia in Cultured Somatic Cells. (A Commonwealth Fund Book.) By John W. Littlefield. Pp. xi+163. (Harvard University: Cambridge, Massachusetts and London, 1976.) £6.15.

THOSE of us who work in the field that is now somewhat pretentiously called somatic cell genetics owe John Littlefield a great deal. This little book puts us further in his debt. It is based on a course taught by him at Harvard over a number of years. The students who were exposed to this teaching can count themselves fortunate. The information presented is constantly subjected to the sort of critical assessment that can be provided only by someone who has himself worked long and intelligently in the field.

As its title indicates, the book has three main themes. The chapters dealing with genetic and other forms of variation in somatic cells and those dealing with the phenomenon of 'senescence' *in vitro* are difficult to flaw. The discussion of senescence, in particular, is informed by a degree of perception that I have not found

matched in any other published account of this subject.

I hope I may be forgiven for saying that I found the treatment of neoplasia less good. I think this may be because the author has not himself had so much experience of work with tumours; whereas, of course, he has made monumental contributions to the study of cells *in vitro*. In discussing experiments on somatic cells growing *in vitro*, he is very conscious of the quantitative aspects of the work he is describing, and especially of quantitative deficiencies in the evidence presented. His assessment of experiments dealing with the formation of tumours is, however, less sure If he had treated the relationship between 'transformation' and tumorigenicity with the same rigour as characterises his analysis of mutagenesis in somatic cells, I doubt whether he would have written (Established) "heteroploid lines appear to have lost growth control and to have become neoplastic". Even so, there's no chapter in this book that isn't worth reading, and there is, as a very important bonus, an absolutely splendid bibliography. **Henry Harris**

Henry Harris is Professor of Pathology at the University of Oxford, UK.

obituary

The recent premature and unexpected death of **James Olds** deprives physiological psychology of one of its most distinguished workers and the neurosciences of a man whose experimental work cannot be ignored in any serious behavioural theory. His major discovery provoked controversy from the time of its first report and although he made substantial contributions himself, many related issues are unresolved at the time of his death.

In 1953 Olds and P. M. Milner, working in D. O. Hebb's laboratory at McGill reported that rats with electrodes in lateral hypothalamic and septal sites would press a lever to deliver trains of stimuli to their own brains. Electrical self-stimulation revealed the existence of powerful central reward mechanisms and immediately raised difficulties for theories in which behaviour is initiated only to mitigate the aversive consequences of internal disequilibrium. The phenomenon had not been predicted by theories current at the time (but the aversive effects of brain stimulation had been looked for and found the previous year by Delgado, Roberts and Miller) and although he had considered the existence of central reward processes, Olds was working on the 'arousing' effects of electrical stimulation when he made his discovery. Noting a tendency for one rat to return to a point where he had previously been stimulated, Olds and Milner were quick

to exploit this phenomenon with an operant conditioning apparatus. Many researchers must previously have seen sniffing and approach behaviours following stimulation through implanted electrodes. The originality of the discovery lay in seeing the possible significance of this behaviour and examining it in a Skinner box.

In succeeding years Olds made major contributions to understanding self-stimulation behaviour. With Travis and Schwing, and later with M. E. Olds, he prepared what are still the most extensive maps of approach and avoidance areas in the rat brain. With Travis, Yuwiler and others he initiated investigations of the possible neurochemical basis of the phenomenon. His suggestion in 1960 that the major tranquillizers act specifically on the mechanisms of approach behaviour has recently attracted new interest in the light of the action of these drugs on dopaminergic mechanisms. Increasingly Olds became interested in single cell recordings as a method of investigation, but he never felt that these had given the answer he really wanted, the anatomical indentity of the reward mechanism itself.

Old's writings on self-stimulation included many provocative theories. He took an intense interest in new ideas and findings which might be relevant to the problem, and in recent years had been preoccupied with the possible role of monoamine neurones in reinforce-

ment. To the end of his life he remained at the centre of controversies concerning the functional significance of this behaviour and its neuroanatomical basis, and in retrospect it was perhaps fortunate that so much work in this field was gathered together at the Beerse conference in 1975 (*Brain Stimulation Reward*, edited by A. Wauquier and E. T. Rolls, Elsevier). James Olds must have impressed many of those present as the man best able to draw together the various strands of evidence in a coherent synthesis. Particular interest attaches therefore not only to his contributions to this symposium but also to the as yet unpublished monograph on Drives and Reinforcements completed a year before his death.

Olds modestly maintained that if he had not discovered electrical self-stimulation it would soon have been uncovered by someone else. This may be doubted, but even if true, few workers could have contributed his enthusiasm and imagination to developing a whole new field of psychophysiological work. He was a man who was willing to consider unusual concepts and paradoxical ideas but pursued them with great clarity. Everyone in the field will feel the loss of his critical intelligence and many will also acknowledge a debt of gratitude for his personal encouragement of their own efforts.

T. J. Crow

announcements

Appointments

Dr H. S. Bedson to the Chair of Medical Virology at the University of Birmingham.

Dr Leo A. Kaprio has been reappointed as Regional Director for Europe of the WHO.

Mr E. S. Booth, chairman of the Yorkshire Electricity Board, as the next President of the IEE.

Professor A. Lazenby, at present Vice-Chancellor of the University of New England, NSW, as the next Director of the Grassland Institute.

Professor D. R. Stranks, of the University of Melbourne, as Vice-Chancellor of the University of Adelaide.

Professor E. W. J. Mitchell, Head of the Department of Physics, as Deputy Vice Chancellor of the University of Reading.

Dr Derek Ogston, Reader in Medicine as Regius Professor of Physiology at the University of Aberdeen.

Dr Roger Grice, of Cambridge University, as Professor of Physical Chemistry at Manchester University.

Professor John Fraser Scott, of the University of Sussex, has been appointed Vice-Chancellor of La Trobe University, Melbourne.

Awards

The Faraday Division of the Chemical Society has awarded the 1976 Marlowe

Medal to **Dr James J. Burton** of the Exxon Research and Engineering Company for his theoretical work on microclusters.

The Finsen awards for 1976 have been made to **S. B. Hendricks** (Control of plant development by light), **H. F. Blum** (Skin and Cancer Hospital, Philadelphia, USA) (Photodynamic action and carcinogenesis by UV light), and **D. Shugar** (University of Warsaw, Poland) (Photochemistry and the structure of nucleic acids and proteins).

The SRC has awarded Senior Fellowships to **Dr A. Boksenberg,** of the Department of Physics, University College, London; **Professor A. Car-**

rington of the Department of Chemistry, University of Southampton; **Professor B. C. Clarke** of the Department of Genetics, University of Nottingham; **Dr P. H. Gaskell** of Pilkington Bros Ltd and the University of Cambridge; **Professor C. A. R. Hoare** of Queens University, Belfast; **Professor E. C. Zeeman** of the Department of Mathematics, Warwick University.

Meetings

October 11–15, **Clean Air**, Edinburgh (The Press and Information Officer, National Society for Clean Air, 136 North Street, Brighton BN1 1RG).

November 4–5, **Mutilation of Animals**, London (Symposium Organising Secretary, RSPCA Headquarters, Causeway, Horsham, Sussex RH12 1HG, UK).

November 8–10, **Manufacture of Superconducting Materials**, Port Chester, New York (Conference Coordinator, Technical Divisions and Activities, American Society for Metals, Metals Park, Ohio 44073).

November 17–18, **Agricultural Yield and Efficiency**, London (The Executive Secretary, The Royal Society, 6 Carlton House Terrace, London SW1Y 5AG, UK).

December 1–2, **Food Additives in the Balance**, Harrogate (Dr D. Howling, c/o Technical Division, CPC (United Kingdom) Ltd, Trafford Park Road, Trafford Park, Manchester M17 1PA, UK).

January 5–7, **Solid State Physics**, Manchester (Deadline for contributions: November 1) (Professor H. E. Hall, Department of Physics, University of Manchester, Manchester M13 9PL, UK).

January 18–19, **Optical Production Technology**, London (Mrs R. G. Keiller, SIRA Institute Ltd, South Hill, Chislehurst, Kent BR7 5EH, UK).

January 25, **Energy Transfer Mechanisms, Dosimetry, Dose-Response Relationships**, London (Professor J. H. Martin, Department of Medical Biophysics, Blackness Laboratory, University of Dundee, Dundee, UK).

February 16–18, **4th International Congress on Mutations: Biology and Society**, Paris (Congrès Service, 1, rue Jules Lefèbvre, F 75009, Paris, France)

March 17–18, **Perinatal Care**, New York (Dr Raymond L. Vande Wiele, Co-chairman of Symposium on Future Directions in Perinatal Care, College

of Physicians and Surgeons of Columbia University, Department of Obstetrics and Gynecology, 630 West 168th St, New York, N.Y. 10032).

March 29–31, **Neuromuscular Disorders**, Bath (Dr G. G. Lune, Department of Biochemistry, University of Bath, Bath BA2 7AY, UK).

April 17–30, **Chemically Induced Magnetic Polarisation**, Sogesta, Urbino, Italy (Professor L. T. Muus, Chemistry Institute, 140 Langelandsgade, DK-8000 Aarhus C, Denmark)

April 23–28, **Annual Meeting of the American Ceramic Society**, Chicago (Deadline for abstracts: November 20) Dr D. N. Winslow, School of Civil Engineering, Purdue University, West Lafayette, Indiana 47907).

April 25–29, **11th International Sym-**

posium on the **Remote Sensing of the Environment**, Ann Arbor, Michigan (Dr Jerald J. Cook, Environmental Research Institute of Michigan, PO Box 618, Ann Arbor, Michigan 48107).

May 1–7, **Stereochemistry**, Burgenstock, Switzerland (Professor P. Pino, Laboratorium fur Technische Chemie, ETH, Universitatsstr 16, CH-8092, Zurich, Switzerland).

May 23–27, **Safety at Sea**, London (The Conference Secretary, WEMT 1977, 10 Upper Belgrave Street, London SW1X 8BQ, UK).

August 29–September 1, **Drugs Affecting Lipid Metabolism**, Philadelphia (Deadline for abstracts: April 15) (Dr William L. Holmes, Scientific Secretary, Symposium on Drugs Affecting Lipid Metabolism, Lankenau Hospital, Philadelphia, Pennsylvania 19151).

August 30–September 2, **Biological Activity and Chemical Structure**, Noordwijkerhout, the Netherlands (Secretariat Department, c/o Merck Sharpe and Dohme B.V., Waarderweg 39, PO Box 581, Haarlem, the Netherlands).

August 31–September 2, **Astronomy and Astrophysics**, Christchurch, New Zealand (Astrophysics Conference Secretary, Department of Physics, University of Canterbury, Private Bag, Christchurch, New Zealand).

September 4–9, **13th International Symposium on Free Radicals**, Lyndhurst, Hampshire, UK (Professor A. Carrington, Chemistry Department, The University, Southampton SO9 5NH, UK).

September 11–16, **Pathology**, London (Professor A. Munro Neville, VIth European Congress of Pathology, Institute of Cancer Research, Fulham Road, London SW3 6JB, UK).

September 12–14, **Electron Microscopy and Analysis**, Glasgow (The Meetings Officer, Institute of Physics, 47 Belgrave Square, London SW1X 8QX, UK).

September 27–29, **Power Semiconductors and their Applications**, London (Deadline for abstracts: November 1) (Press Officer, IEE, Savoy Place, London WC2R 0BL, UK).

October 24–26, 1977, **Erosion: Prevention and Useful Applications**, Vail, Colorado (Deadline for abstracts: March 30) (Dr William F Adler, Effects Technology, Inc., 5383 Hollister Avenue, Santa Barbara, California 93111)

October 21, 1976

The price of trouble: £300

SHOULD university students be offered differing sizes of government grants depending on the subject they choose to take and its perceived relevance to national needs? The question has come into sharp focus in the United Kingdom with persistent rumours that the Department of Industry is trying to get engineering students paid an extra £300 on their annual grant. At present most students with state support get nominally identical grants, though some who receive grants from industry or the armed forces and who may have commitments after leaving university receive higher sums. Differential grants have been used in the past—in the 1930s potential teachers were given a supplement as undergraduates.

There are three questions which must be asked of any such scheme—is it needed; will it be effective; will it generate unnecessary trouble? There could be a need for differential grants if engineering graduates were in high demand, if universities had room for students and yet they were unforthcoming in sufficient quality and quantity. Obviously a simple label 'engineer' is inadequate for determining who is in high demand, and equally obviously many businesses which say they don't need more graduates are deceiving themselves. Even so, the supply of graduate engineers, whatever its defects, does not seem to have rung too many alarm bells. Nor, contrary to public opinion, are university engineering departments lonely places to be in these days. In 1976 there are 36% more students starting civil, mechanical and electrical engineering courses than there were in 1973 (during the same period the entry into mathematics, physics and chemistry dropped by 4%). The school academic attainments of engineering undergraduates are not conspicuously good, but it is by no means clear that £300 will buy quality; which brings us to the question of effectiveness of the measure.

If the measure turns out to be effective, it could easily be so by diverting students from adjacent disciplines: to collect £300 mathematicians would become mechanical engineers; physicists, electrical engineers; chemists, chemical engineers. Some effectiveness! Further, the measure can only be effective if schools are capable of responding to changing demands and are capable of giving wise advice to their pupils on whether

the money justifies the switch in individual cases—yet many believe it is in schools where counselling on engineering careers is inadequately done. And the sort of student who will be tempted by £300 is also likely to look at the way engineering salaries compare with those in other professions—and maybe to decide that how the Department of Industry and how industry itself rates him are two different things.

Finally, the troubles it will cause. If engineering, why not science, for technology uses research results. Why not economics, because it is an economic mess we are in. Why not business management, because our companies need to be run better. Why not agriculture, because food imports make a big hole in our balance of payments. Why not design, which sells commodities. Or art, music, drama, which bring in tourists. The idea once accepted in one instance could lead to an almost endless queue of special cases. And once the scheme is started, how is it stopped without undesirable effects? After all, when ICI cut its graduate recruitment of chemists one year the consequences reverberated for years in schools and universities. So when the government phases out its supplement there can be no doubt that it will depress engineering recruits.

So, for a variety of reasons, the £300 scheme seems to be all wrong. But that should not deflect us from a deep concern that engineering does need some sort of leg-up from its present low-prestige intellectually-unsatisfying status. One thing about which the Department of Indsutry should be talking seriously to the Department of Education and Science is whether schools can provide adequate support and encouragement to potential engineers. A second is whether too many universities don't impose rigid barriers between mathematics, science and engineering thereby preventing students from 'tasting' engineering without making an irrevocable commitment.

It is not clear that opposition to differential grants at undergraduate levels should necessarily mean opposition at the postgraduate level, where the sacrifice of an industrial salary for the spartan life of a higher degree may for many engineers be intolerable in a way that it is not for most other research students who lack outside attractions. □

A view from home

Andrew M. Sessler, Director of the Lawrence Berkeley Laboratory (LBL) at the University of California, Berkeley, responds here to the assessment of LBL carried by Nature *(August 12, page 528)*

THE article on LBL was, in my opinion, neither fair nor accurate. I acknowledge the right of a writer to put down what he believes he has observed, but I cannot in good conscience say that I recognise our Laboratory either in its past or present state. In describing its past, for example, the writer, Wil Lepkowski, quotes Robert Yaes on the failures of Lawrence's Laboratory to discover artificial radioactivity and nuclear fission. Those are true charges, but they should surely be considered against a background of the many revolutionary discoveries which were made in the 1930s when Lawrence's Laboratory led the world in nuclear science.

Many fundamental features of nuclear reactions, nuclear structure, radioactive decay, nuclear chemistry and artificial radioactivity were established. The Laboratory was the birthplace of the transuranium elements, and has active research currently in progress to search for possible superheavy elements. The Laboratory established the science of nuclear medicine before the Second World War, when most of the radioactive tracers employed in medicine today were discovered and the foundations were set for the tracer techniques and scintillation detector cameras that are now so commonly used in diagnostic medicine. LBL's eight Nobel Prizes in fact reflect its vitality in basic science throughout Lawrence's time.

Mr Lepkowski is correct in his point that LBL no longer dominates the world of high energy physics, but the sense of failure he attaches to this seems unwarranted. It would be absurd to expect that any laboratory could continue to dominate such a thriving field. Yet our high energy physicists from LBL are not in fact doing all that badly. When the *Nature* article appeared the high energy physics community throughout the world was excitedly discussing the first observation of a charmed particle by a long-standing Lawrence Berkeley Laboratory-Stanford Linear Accelerator collaborative group. This same collaboration had made a major discovery two years earlier of the ψ/J particle, at the same time that the same particle was discovered at Brookhaven, and subsequently they discovered the ψ'. Furthermore, the collaboration of LBL and SLAC is a joint project in every

sense of the word: they have been working together on this project for more than five years, using a national facility supported by the entire high energy physics community in the United States. LBL is proud to be an essential part of its design, construction and management, and can expect to continue to have a significant role in high energy physics research for decades to come.

Mr Lepkowski overlooks the strong and unique position that LBL occupies in heavy-ion nuclear physics. Our Laboratory has major programmes at its sector-focusing 88-inch cyclotron, the SuperHILAC and the Bevalac, the latter being the only source of relativistic heavy ions in the world. LBL accelerators are operated as national facilities and there are strong in-house experimental and theoretical groups. We are not blind to the fact that formidable competition is growing, particularly in Europe which is now allocating far more support to nuclear physics than is the United States, but at this time LBL is still considered by all to be a major international centre for nuclear physics.

Basic research is also very healthy at LBL in fields too numerous to catalogue in detail. Fundamental studies in biology and medicine work range from yeast genetics, physiology, and lipoprotein studies to radiobiology and nuclear medicine. LBL, in collaboration with others, is currently pioneering clinical trials on the efficiency of relativistic-energy heavy ions in cancer therapy. In molecular science we have some of the world's forefront studies on reactions in crossed molecular beams, on laser induced chemistry, and on quantum electrodynamic studies in high-Z hydrogen-like atoms.

In material science surface polaritons and plasmons are being studied by means of their nonlinear excitation through optical mixing of laser beams; and synchroton radiation is being employed, along with low energy electron diffraction (LEED), to ascertain the nature of catalytic surfaces. The fundamental understanding of alloys by use of electron microscopy and other advanced techniques has led to the ability to design steels and alloys of unprecedented toughness and hardness. From the field of astronomy there is the first measurment of the mass of

a neutron star. In sum, basic science at LBL is diverse, excellent and enthusiastically pursued.

Mr Lepkowski dwells negatively on changes going on in LBL to address changing national and societal needs and a new set of goals from its major funding agency, the US Energy Research and Development Administration. Adaptation to a changing world is a necessity for organisations as it is for individuals. Furthermore, it is simply false to state that the old leaders and, in particular, the Nobelists, are still the dominating leaders of the Laboratory. In fact the Director, the Deputy Director, and eight of the ten Associate Directors have held their positions for less than four years. We of the present management make no apologies for changing the character at LBL.

We can also point to numerous entries into new fields in which important contributions have already been made and which have exciting prospects for the future. Among these are geoscience, geothermal resource assessment, controlled thermonuclear power, metallurgy, catalysis, isotope separation, materials science, environmental and biological effects of energy technology pollutants, energy policy analysis, solar power, and energy conservation studies. In our experience, when we enter fields which are non-traditional for LBL the quality of our scientific staff is revealed by immediate innovative use of advanced experimental techniques and theoretical analysis which result in important or revolutionary contributions to the new fields.

I can give the example of photo-electron spectroscopy turned to the study of the role of power plant particulates in the catalytic conversion of SO_2 to (harmful) sulphates; or the development, by a scientist previously working on the Lamb Shift, of the most sensitive detector for heavy metals by Isotope Shifted Zeeman Atomic Absorption (IZAA); or the development of superconducting quantum interference devices (SQUID) to a sensitivity some orders of magnitude greater than that available elsewhere, and their direct application to geothermal resource exploration. Or I could cite our work, some years ago, which initiated concern over the catalytic destruction of the protective ozone layer by supersonic transports.

I could give several examples, but let it suffice to say that we feel comfortable in our new activities because we know we are doing excellent science and engineering on problems of importance to mankind. ☐

Exotic viruses

Arie Zuckerman offers this brief on some
of the diseases now attracting broad attention

A BRIEF perusal of the daily press over the past few months would reveal an incidence of unusual disease which, to many, seems quite remarkable. There have, most recently, been the illnesses reported from the Sudan, Nigeria, Zaire and Sierra Leone— now apparently identified, by the World Health Organisation last week, as Marburg disease. Work on the identification of Philadelphia's so-called Legion's Disease continues in the United States. Lassa fever meanwhile attracts considerable attention. Rabies and Swine influenza receive even greater emphasis. Even Bubonic Plague has been in the headlines.

Of the many viruses that can infect more than one species of animal, three are of particular current interest. Rabies, with a huge reservoir of infection in the wildlife population of most parts of the world, is now racing across Europe towards Britain, where it has not been endemic for more than half a century. Marburg disease, first identified nine years ago, is still not fully understood. And Lassa fever looms as an awesome viral infection in tropical Africa.

Marburg disease

Marburg virus (or "green monkey") disease, a severe, distinctive, febrile human illness, was first described in 1967. Thirty-one cases of illness with seven deaths in Germany and Yugoslavia were traced to direct contact with blood, organs or tissue cell cultures from a batch of African green monkeys caught in Uganda. Several secondary cases subsequently occurred in hospital personnel who had been in contact with blood from patients; one further case was apparently transmitted by sexual intercourse. Although 29% of the primary cases proved fatal, there were no deaths in the 6 secondary cases.

It soon became apparent that this was a previously unrecognised disease, caused by an infectious agent probably new to medical science. Laboratory work using guinea pigs and Rhesus and vervet monkeys revealed an elongated sinuous organism generally curled in the form of a figure 6 or a horseshoe with superficial resemblance to rabies virus. But Marburg virus is morphologically distinct and does not share any antigenic properties with rhabdoviruses or with any other known viruses.

The infection has been investigated

intensively since the first outbreak, but the natural cycle of transmission in nonhuman primates and perhaps other species of animals and the origin of the infection remain unknown. The disease has not hitherto been recognised in a sporadic form in Africa.

The first recognised outbreak of the disease in Africa, and the first since the original outbreak, occurred in South Africa in February, 1975. The victim, a young Australian man who had hitch-hiked through Rhodesia, died in a Johannesburg hospital; shortly afterwards his travelling companion and one of the nurses who had cared for him contracted the same disease. Both recovered. Virological investigations showed the outbreak was caused by the Marburg virus. As in the original outbreak in Marburg, when the virus was isolated from semen 83 days after the onset of illness evidence was obtained that the virus could persist in the body for at least 2–3 months after the initial attack; the virus in this case was isolated from the victim's eye fluid 80 days after the onset of the illness.

Lassa fever

The dramatic story of Lassa fever began when two missionary nurses from Lassa, in North-Eastern Nigeria, died in 1969 from a mysterious illness; a third, who was gravely ill and flown to the United States, recovered. Convalescent plasma from her was effective in the treatment of a laboratory worker, who acquired the infection while working with tissue cultures infected with blood from these patients.

The following year a further outbreak of Lassa fever killed 52% of 23 patients admitted to hospital in Nigeria. Dr Jeannette Troup, who was primarily responsible for drawing attention to this disease, also died of the infection. Since then further cases have occurred and 10 out of 21 medical workers have died from Lassa fever. Up to June, 1975 the disease was recognised on nine different occasions involving 114 officially recorded cases, two of which resulted from laboratory infections. Over a third of these infections were acquired by person-to-person spread within hospitals. Many other clinically suspect cases have since been uncovered from reviews of hospital records: retrospective searches through the medical literature have revealed references to cases closely resembling Lassa fever in the Central African Republic, Upper Volta, Nigeria and Sierra Leone. Since 1969, cases have occurred in Nigeria, Sierra Leone and Liberia, and serological evidence suggests the virus may have been active elsewhere in West and Central Africa. And of course recent cases identified in air travellers to Britain, the United States and Canada have received extensive press coverage.

The clinical spectrum of Lassa fever extends from very mild or inapparent illness to fulminating fatal infection. Among patients admitted to hospital the death rate is as high as 67%, with more fatalities among Caucasians than Asians. Intimate contact with any primary case appears to carry a high risk of infection; relatives or medical attendants providing direct personal or nursing care are particularly threatened. Cases have occurred among patients and hospital visitors who apparently had no direct contact with the primary cases. Only a few such tertiary infections have occurred, and the reasons for the interruption of
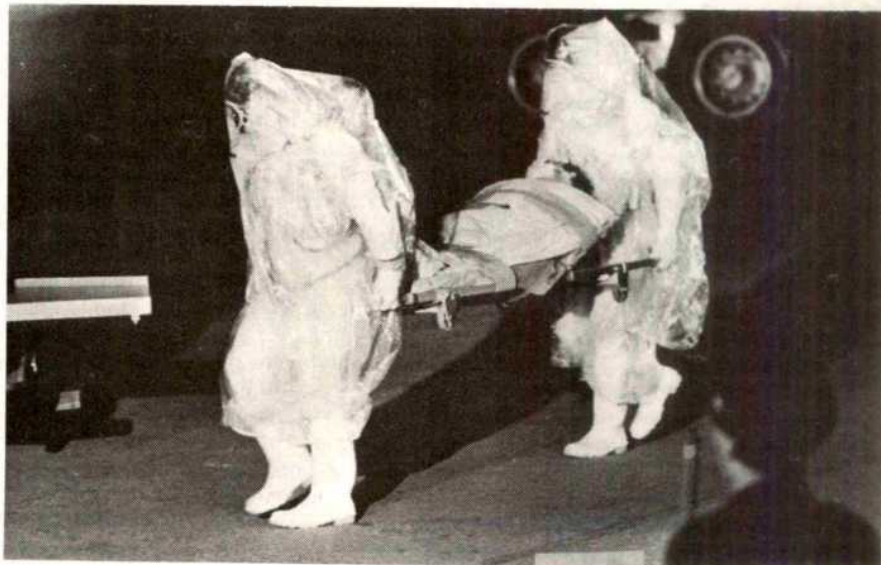
Lassa patient arrives in Hamburg from Nigeria

Keystone

outbreaks after the generation of secondary cases are not clear. It seems that, although the primary cases were highly infectious, the secondary cases were not.

Lassa virus is a member of the arenavirus genus, a group of enveloped RNA viruses with a virion of granules instead of a defined core. Along with Machupo virus (which causes Bolivian haemorrhagic fever) 'it 'produces a significant mortality in man. Lassa appears to have the greatest potential for direct spreading in human populations. It is related antigenically to the rodent-associated haemorrhagic viruses of South America and to lymphocytic choriomeningitis virus, a persistent infection of mice. It has been isolated from a single rodent species, the multimammate rat *Mastomys natalensis* in West Africa.

More recently, a virus closely related immunologically to Lassa virus was recovered from the tissues and organs of wild *Mastomys natalensis* captured in 1972 in Mozambique. This finding more than doubles the potential area of Lassa-like virus distribution in Africa The multimammate rat is the most widespread and commonest rodent in Africa south of the Sahara, and an ideal carrier to the domestic environment because of its semi-commensal habit and a combination of other factors, including an exceptionally high propagation rate. Whether all the *Mastomys* rodents are capable of maintaining and transmitting Lassa virus, is not known, and the ecological interrelationships between *Mastomys* and the presumed non-reservoir genera *Rattus* and *Mus* have yet to be determined. Carnivores, in particular dogs and cats which probably prey upon *Mastomys* and other rodents, have not yet been investigated as potential secondary hosts.

The Lassa virus may spill over from the rodent cycle to man by various routes including, for example, contamination with rodent urine, directly or on foodstuffs or through dust. The means by which the virus is spread from one person to another is not clear. Accidental inoculation with a sharp instrument or contact with blood has accounted for a few cases. The virus has been isolated from the pharynx and urine of patients, so that indirect airborne spread of the virus as well as mechanical transmission are most likely.

Lassa fever is not a disease which comes under international quarantine regulations, which at present offer the only effective means of reducing the risk of transmission. Transmission can be limited to some extent by surveillance, identification, and the hospitalisation and isolation of patients, but quarantine of households may have

to be considered by public health authorities, and ways of reducing rodent populations should be explored. The evacuation of patients presents special problems. Until effective methods of treatment and control become available and a suitable vaccine is developed, the epidemiological propensities of this infection may be very wide with the modern means of air travel.

Rabies

Rabies virus is a member of the rhabdovirus group, which includes a diverse collection of characteristic bullet-shaped RNA viruses from mammals, reptiles, fish, insects and plants. The concept of a single antigenic entity is no longer accepted since there are four recognised rabies-related viruses, although these appear to be confined to sub-Saharan Africa. Two of these viruses, Mokola and Duvenhage, are associated with fatal rabies-like human illness.

Rabies is a viral infection of the central nervous system which has been recorded in most domesticated and wild warm-blooded animals. The virus, present in salivary secretions, is transmitted mainly by bites or scratches, although the inhalation of minute droplets in the air can also lead to infection, and oral transmission has been demonstrated in experimental animals. The susceptibility of an animal depends on its age, the degree of infection, the site of introduction and the properties of the particular virus strain involved. Once the virus leaves the site of inoculation it usually passes to the central nervous system through the peripheral nerves. Once the · central nervous system has become infected the virus may multiply, attacking such tissues as the salivary and other glands, the kidneys, the brown fat, and, in some species, muscle and lung. In some infected organs the virus is found primarily within the nerves.

The effects of the disease on the nervous system and the change of behaviour in the victim are probably important to the mode of transmission. Typically, the sick animal (and this applies particularly to small carnivores) becomes unusually and indiscriminately aggresive, while in the earlier stages of infection timid animals become less restrained, approachable and friendly High population density and turnover and long incubation periods maintain the infection even though it is usually fatal Man is irrelevant to survival of the virus: there is no adequately documented case of human transmission of rabies, although a theoretical risk exists because the virus is present in secretions

The most important pool of infec-

tion in most areas of the world is the wildlife population, including, in particular, foxes. Examination of thousands of small wild rodents in the Americas has not revealed rabies infection, but in parts of Europe the isolation of rabies virus with unusually low virulence has been reported from wild rodents. The significance of this finding is not clear, and cannot be immediately associated with fox rabies. The ecology of bat rabies is intriguing· in the Americas, numerous species of bats have been found with rabies infections, but in many areas these infections are independent of other mammalian cycles. The disease is readily passed on to cattle by the bite of vampire bats during feeding; in Latin America, for example, bat rabies is estimated to have killed many hundreds of thousands of cattle. Over 170 human deaths from rabies have been attributed to vampire bats, which suck the blood of sleeping victims. But natural transmission from insectiverous bats to other terrestrial animals by biting has not been observed, and experimental transmission with infectious saliva has proved difficult.

The major problem caused by rabies, according to a paper from the Office of Health Economics in the UK, does not stem from its immediate effects, but from the medical, social and economic risks and the cost associated with its prevention Improved human diploid cell vaccines are in fact now becoming available for pre- and post-exposure prophylaxis in man, while supplies of human rabies immunoglobulin for passive immunisation and local wound treatment are expected to increase and developments in intensive care promise a better outlook for future victims.

The control of rabies in animals, on the other hand, presents considerable difficulties The international transfer of animals must be rigidly controlled. and strict precautions should be taken in the acquisition of mammals from the wild. The control of wildlife rabies is even more difficult Hormonal sterilisation can be effective experimentally but has not yet been used successfully in the field. The immunisation of wildlife may become feasible in the future, particularly oral immunisation of foxes using bait doped with live attenuated virus. The control of bat rabies requires further investigation. although inoculation of cattle with small amounts of anticoagulant has proved lethal for vampires. Similarly the backs of captured vampire bats can be smeared with anticoagulant mixed with petroleum jelly: their release has resulted in dramatic reduction in the number of bats, since the treated bats contaminated others in their colony by contact.   □

NUCLEAR TRADE

More words than action

International trade in nuclear technology was a key issue at the 20th annual meeting of the International Atomic Energy Agency (IAEA) held in Rio de Janeiro last month. Bruce Handler reports from Brazil

ONE reason the IAEA meeting was held in Latin America was to encourage a better understanding of the problems facing developing countries in the field of nuclear energy. Nearly 1,400 delegates who attended wound up hearing almost as much about problems IAEA members have with each other as they have with nuclear energy, because a good portion of the debate consisted of political tirades. But everyone was in favour of the further development of nuclear energy. The sharp divisions were over which countries should assume the responsibility for developing it, and to what extent.

Many industrialised nations expressed fears over putting sophisticated nuclear technology into the hands of underdeveloped Third World countries, principally because of the widely shared belief that such countries could be encouraged or provoked into diverting this knowhow for military purposes. A corollary fear was that terrorists could stage a raid on an underdeveloped country that possessed nuclear reactors and steal the necessary components for making an atom bomb, even if the country in which this occurred was actually using atomic energy only for peaceful purposes.

The Third World rebuttal of this position was that underdeveloped countries could never progress if they are not allowed to acquire modern technology, in nuclear energy as well as in other fields. Developing countries described as discrimination the desire on the part of industrialised nations to limit the spread of such technology.

The IAEA's director-general, Sigvard Eklund, noted that many Third World countries have not signed the 1970 Nuclear Non-Proliferation Treaty because of this feeling. He appealed to them to reconsider the political reasons which originally led to their taking such a position but later recognised, realistically, that when it comes to national pride, nations sometimes refuse to listen to the most soberly reasoned arguments.

The country most strongly opposed to the sharing of nuclear technology was the United States. The head of the US delegation, Robert C. Seamans Jr., director of the US Energy Research and Development Administration, said in an address that the USA did not want to see sophisticated atomic knowhow spread throughout the developing world.

He said that countries which are taught how to enrich their own uranium for use in nuclear reactors could use this knowhow to separate plutonium from uranium. This, in turn, Seamans stated, "can increase the risk of diversion (of fissionable material) to nuclear weapons and also the risk of terrorist activities." The United States was willing to enrich uranium for other countries and make it widely available, but it was a risk to international security to show developing nations how to do their own enriching. Seamans added that as a compromise, the IAEA could set up regional uranium-enriching centres, which would at least be under multinational control.

Yugoslavia presented a proposal for creating a multinational pool of nuclear resources, to make it easier for developing countries to begin or expand nuclear energy programmes.

The IAEA meeting, however, was basically a forum, rather than a policy-making session, and the best any single delegate could hope for was that others would consider the proposals discussed and take them back to their home governments for study and possible implementation.

The Third World view on technology sharing was expressed most emphatically by Hervasio de Carvalho, the president of Brazil's Nuclear Energy Commission. He labelled the preoccupation over safeguards and the spread of nuclear technology to underdeveloped nations as "exaggerated", and repeated what Brazilian government officials have been trying to

---

One complete session of the IAEA meeting was devoted to the scientific discussion of advances that have been made in the use of nuclear technology in food and agriculture. Among the highlights:

● Dr Maurice I. Fried, of the United States, director of the Joint FAO/IAEA Division of Atomic Energy in Food and Agriculture, gave a historical introduction to the use of nuclear techniques in the field of food and agriculture, and told delegates that isotope and radiation techniques "are neither esoteric nor so sophisticated that developing countries cannot apply them for the solution of practical problems." He noted that the bulk of the work of the joint FAO/IAEA programme he heads is aimed at helping developing countries use atomic energy to improve their agricultural output.

● Dr Dick De Zeeuw, of the Netherlands, general director of agricultural research of the Netherlands Ministry of Agriculture and Fisheries, offered the sombre forecast that by the year 2010 the world will have 1,800 million more people who consume less than the required daily amount of food than it does today, and that most of these new hunger victims will be in Southern and Southeast Asia. Many developing countries, he said, now lose from 20–40% of their crops after harvest; food-irradiation and related techniques could reduce this loss, but the main obstacle was the failure of policy-makers in developing countries to adopt the widespread use of atomic irradiation of food and to discard ineffective traditional methods.

● Dr K. Sundaram, of India, director of the biomedical group of the Bhabha Atomic Research Center in Bombay, made the point that effective food preservation techniques are difficult to transfer to underdeveloped ones, because of problems of money, materials, technology and energy of the poor nations' end. He said that developing nations around the world had failed to recognise irradiation's true importance, and criticised them for wrongly copying food preservation methods from developed countries, which, he noted, do not need to place so much importance on irradiation because of the already high standards of their food industries.

● Admar Cervellini, of Brazil, professor and head of the Physics and Meteorology Department of the University of Sao Paulo's Luiz de Queiroz Agricultural College, in Piracicaba, spoke about the work of Brazil's Centre for Nuclear Energy in Agriculture, of which he also is the head. The centre, which was set up with IAEA assistance, is now ten years old.

Dr Cervellini said that one of the centre's most practical achievements has been the development, through radiation mutation, of strains of beans with increased yield, with higher protein content, and with more resistance to Golden Mosaic disease. Seeds of the latter variety have been distributed to five other Latin American countries.

explain to the rest of the world since the controversial nuclear deal with West Germany in 1975: that Brazil, which is poor in oil, has low-grade coal and is quickly using up its feasibly exploitable hydroelectric potential, has no other alternative but atomic energy for meeting its future electric-power needs. And for reasons of national security and common sense, Brazil cannot remain dependent on outside suppliers for nuclear fuel—as the United States would like.

Carvalho pointed out that even though Brazil has not signed the Nuclear Nonproliferation Treaty (for the usual reason: opposition to the division of the world into an elite, industrialised "atomic club" and everyone else), it has submitted the agreement with West Germany to IAEA safeguard supervision. This, he said, is proof that Brazil is not going to make any bombs, adding that the world should accept the word of other countries that make similar pledges to use atomic energy only for peaceful purposes. (The IAEA Board of Governors, which met in Rio before the general meeting, approved safeguard provisions in two other major transfer-of-tech-

nology nuclear agreements: the one between Canada and Spain and that between France and South Africa.)

The Nigerian delegate's contention, that developing countries are in the IAEA "to be seen but not heard" and that they are "merely tolerated" by industrialised countries, did not seem to be reflected by the vote on a proposal sponsored by Iraq to admit the Palestine Liberation Organisation to the IAEA, as an observer. Of the 71 of the 110 member countries of the agency that were present for the vote, 46 voted in favour, 21 abstained and only four (Israel, the United States, South Africa and Costa Rica) voted no.

A move led by black African countries and supported by several socialist states to expel South Africa from the IAEA, on grounds that the Pretoria government does not represent the people, did not succeed. The best these countries could do was to get a motion passed to review South Africa's permanent seat on the Board of Governors, but with no action to be taken until the 1977 IAEA meeting. It is unlikely that even this watered-down proposal will bring results, because IAEA by-laws say the permanent

regional seats on the board must go to the countries in each specific area of the world that are the most advanced in nuclear technology. In Africa, there is no doubt that this is South Africa.

The effect of nuclear power stations on the environment was another topic to come up at the meeting. Eklund expressed veiled dislike for citizens' movements which have sprung up in countries such as the United States to stop further construction of nuclear reactors. The IAEA chief asserted that nuclear power plants are "friends of the environment." He said there has never been a fatal accident from nuclear causes at any of the 180 non-military nuclear installations now operating around the world. In contrast, Eklund noted, there have been serious accidents at hydroelectric plants, including the bursting of dams and the consequent widespread destruction of the environment.

Debate continued at the meeting on how best to dispose of waste material from nuclear reactors, by burying it in the ground, storing it in protected containers, dispersing it into the air or dumping it into the ocean. Perhaps predictably, no consensus was reached.

USA

New directions in lobbying

A new lobbying organisation concerned with global scientific environmental and social issues was established in the United States last week. Colin Norman reports from Washington

ALTHOUGH Washington is thick with lobbyists, public interest groups, research organisations and assorted political hucksters selling commodities ranging from the B-1 bomber to environmental awareness, the debut last week of a new 'citizens lobby' deserves —and is getting—more than passing attention.

Called 'New Directions', the organisation will be devoted to raising Washington's political consciousness (or at least, the consciousness of Washington's politicians) on such international issues as arms control, population growth, environmental pollution, and the depletion of resources. It will be an unabashed lobbying enterprise, focusing its attentions on the Congress. The reason it is already attracting considerable notice is that the list of its officers and governing board reads like a *Who's Who* of people concerned with such matters.

Full-time President of New Directions is Dr Russell W. Peterson, former

Governor of Delaware, who recently resigned as Chairman of the Council on Environmental Quality to head the new group. And the outfit's executive council will be chaired by Dr Margaret Mead, the well-known anthropologist who has long been active in public interest causes associated with science.

Planning for New Directions began in the Summer of 1974, when Norman Cousins, editor of *Saturday Review*, and Theodore M. Hesburgh, president of the University of Notre Dame, convened a small group to discuss whether such an organisation is needed and if so, what form it should take. During the past two years, the planning group has expanded to more than 100 people, including such luminaries as Robert McNamara, head of the World Bank, Kingman Brewster, President of Yale University, James Grant, president of the Overseas Development Council, Lester Brown, head of the Worldwatch Institute, and Sargent Shriver, former director of the Peace Corps.

The organisation will be similar in structure to Common Cause, a lobbying group which concentrates its attentions on domestic affairs. It will be a grass-roots organisation, financed by membership fees of $25 per year, solicitations for which will soon be made through a direct mail campaign

using membership lists of sympathetic organisations. New Directions hopes to attract at least 300,000 members, a number which would give it a very sizable war chest to put its message across.

Although the general nature of that message is evident from the concerns of the people associated with New Directions, priorities and details of the organisation's objectives are now being mapped out by groups of individuals working in the general areas of arms control, poverty and the widening economic gap between rich and poor nations, environment and natural resources, human rights, and governmental decision-making.

There are already numerous organisations active in Washington in those areas, but New Directions hopes to provide an additional function—political lobbying to get the concerns raised by various public interest organisations through to the policy makers. Few existing organisations have been able to take on such a role because they would lose considerable tax advantages if they were to indulge in direct lobbying activities. Essentially, they are registered as educational or research outfits. Though a recent change in the tax laws has lifted some of those restrictions, few organisations are geared up to lobby actively.

A statement put out last week by the planning committee of New Directions

describes its relationship with existing groups. "While the work of these groups provides an essential base of information and documentation," it says, "the missing element has been a force unlimited in scope and unfettered by tax prohibitions . . . No organisation has set out to mobilise existing resources to take hard political action on the critical measures which must be implemented to relieve the problems which now disrupt all economies, undermine political stability, suppress human freedoms, and directly threaten human survival throughout the world. Our purpose is to build such an organisation".

The idea is that New Directions will work closely with other groups—indeed, many of them have representatives on the organisation's governing board—and draw heavily on their research and expertise. It will select issues to lobby for, try to bring together informal coalitions of groups to work on them, and serve as a coordinating centre for exchange of relevant information. But its chief role will be to draw as much attention as possible to global problems and their potential solutions, through direct lobbying with members of Congress, government officials, corporate executives and so on. It will also take its case to the news media, organise local groups, go to court when necessary, and provide direct support to candidates for political offices.

Clearly, although the organisation will derive much of its influence from its star-studded cast of backers and officials, the crucial factor in its effectiveness will be the size of its membership. Politicians tend to listen more attentively to messages delivered by groups backed by large numbers of voters or by large amounts of cash which can be used for campaign support. If New Directions can succeed in attracting hundreds of thousands of members, its voice will be heard above the general background noise in Washington.

At this point, however, it is difficult to predict whether or not the organisation can count on that level of support. The people most likely to join are probably already members of existing groups such as Common Cause, and it remains to be seen whether the depth of their concerns extends to spending another $25. But there is at least one promising indication: the planning committee decided to go ahead only after a survey by a New York market research firm found considerable support for the goals of the organisation.

Another potential problem for the organisation is that unless it treads carefully, it may end up stepping on the toes of existing bodies which are already concerned with global problems

FDA's aerosol ban

In a statement which caught most observers here completely by surprise, the Food and Drug Administration (FDA) announced last week that it intends to phase out the use of chlorofluoromethanes (so-called fluorocarbons) as propellants in many types of aerosol sprays. Although the statement did not say when the ban would take effect, or how it would be implemented, FDA said that it would begin by requiring warning labels to be placed on aerosols containing fluorocarbons, and that the ban would be imposed in an "orderly" manner.

The agency's action is surprising since it follows hard on the heels of a major report by a committee of the National Academy of Sciences which recommended that regulation of fluorocarbons should be delayed for up to two years to allow time for more research on the mechanism by which fluorocarbons are believed to be breaking down the ozone layer. The academy report suggested that since the rate of destruction of the ozone layer is small, a two-year delay in regulation would present little additional hazard.

Dr Alexander M. Schmidt, the Commissioner of FDA, said last week, however, that additional research would narrow the range of uncertainty in the calculations of ozone depletion but it "won't change the ultimate regulatory situation". He argued that "given the effects on human health even a 2% ozone depletion from 'unessential' uses of fluorocarbons is undesirable".

"The known fact", Schmidt said, "is that fluorocarbon propellants primarily used to dispense cosmetics are breaking down the ozone layer. Without remedy, the result could be profound adverse impact on our weather and on the incidence of skin cancer in people. It's a simple case of

negligible benefit measured against possible catastrophic risk".

FDA's authority extends over foods, drugs and cosmetics. Thus, in theory, it has the power to regulate the formulation of products such as hair sprays and anti-perspirants which account for the bulk of aerosol uses of fluorocarbons. According to FDA's statement, the agency has authority over about 80% of all products now packaged in aerosol containers.

Schmidt said that FDA will publish details of its proposal to require labelling of aerosols containing fluorocarbons in mid-November, and that details of the phase-out programme will follow a few weeks later. The labelling programme itself should reduce consumption of fluorocarbon-containing products, and it will also help discourage stockpiling.

A spokesman for the DuPont company, the chief manufacturer of fluorocarbons, said last week that he was "astounded" by FDA's decision. The industry had regarded the academy's report as a victory because, in recommending a two-year delay in regulation, the academy had essentially backed the industry's argument that more research is needed to settle scientific uncertainties.

The industry is unlikely to take FDA's action lying down. Although industry spokesmen would not say last week how they expect to contest the ban, it is likely that they will ask for public hearings and, if necessary, go to court to prevent FDA putting its proposals into effect. One possible legal challenge may involve whether or not FDA has the authority under existing laws to take such action. The intensity of the industry's reaction will depend largely on how much time FDA gives it to phase out use of fluorocarbons, however, and that won't be known until towards the end of the year.

—such as the United Nations Association, the Overseas Development Council and the Worldwatch Institute. A check with people from some existing groups last week, however, found little fear that New Directions would steal their thunder or interfere with their goals. In fact, most welcomed the possibility of having a heavyweight group fighting for their causes.

Some supporters of New Directions also argued last week that one of the organisation's chief assets is the fact that it is headed by Russell Peterson. A PhD chemist who worked for DuPont for 26 years, Peterson is a

Republican who built a good reputation during his term as Governor of Delaware between 1969 and 1973. Peterson also knows his way around the Washington power structure very well, having served as chairman of the President's Council on Environmental Quality for three years. During that time, he took several independent stands, the latest of which was to call for immediate regulation of fluorocarbons in aerosol sprays. As one observer put it last week, a Republican, a former governor, and a former top adviser to the President are good credentials for a lobbyist. □

USA

● The International Council of Scientific Unions (ICSU) is the latest organisation to seek a role for itself in the mounting debate over the risks and benefits associated with recombinant DNA experiments. At its 16th General Assembly, held in Washington last week, ICSU decided to establish a Committee on Genetic Experimentation to undertake a variety of tasks, ranging from the provision of advice and information to anyone who seeks it, to the support of lectures and training courses on safety techniques. The committee will also look into the possibility of providing ICSU support for a facility where large populations of cloned fragments of, for example, mouse or human DNA will be constructed and maintained for use by individual scientists. One aim of the committee is to use ICSU's considerable influence to try to ensure that the various controls on recombinant DNA experiments which are being imposed by countries around the world are consistent with each other. As Sir John Kendrew, ICSU's secretary general put it last week, it doesn't make much sense for one country to put very strict controls on the rsearch and for another to allow experiments to go ahead with few restrictions.

But Kendrew acknowledged that the countries with which ICSU deals "are sovereign states and you can't tell them what to do". The committee (which has the snappy acronym COGENE) will thus work behind the scenes, collecting and distributing information and providing assistance to national committees where it can.

Part of COGENE's task will be to ensure that recombinant DNA experiments are allowed to proceed, albeit under "appropriate and generally agreed safeguards". Dr William J. Whelan, who headed a committee which wrote COGENE's charter, noted last week, for example, that "the committee is not coming into being to preside over the demise of genetic manipulation and to that extent, ICSU is expressing a position".

The committee's most valuable function, however, may well be its provision of training courses and other technical assistance for recombinant DNA researchers. Kendrew pointed out that in smaller countries training in biological safety techniques may be limited, and ICSU could provide a needed service. The committee will also look into ways to aid in the distribution of strains of crippled hosts and vectors for recom-

binant DNA work, to help ensure that such strains are universally available.

As for the suggestion that COGENE should provide assistance for a central clone bank for fragments of mammalian DNA, the establishment of such a facility could greatly reduce the proliferation of some of the more hazardous cloning experiments. A similar clone bank is under con-

sideration in the United States, for use by American scientists.

● Once again, the Food and Drug Administration (FDA) has refused to allow the artificial sweetener cyclamate back on the market in the United States because of unanswered questions about its safety. Cyclamates were banned by FDA in 1969 on the basis of studies which indicated that they may cause bladder cancer in test animals when fed to them in large quantities; Abbott Laboratories, the manufacturer of the sweetener, contested the ruling and has been trying ever since to get it reversed. Last year, Abbott presented FDA with a petition, backed by a sheaf of test data, purporting to show that cyclamates are safe. But, last week, after turning Abbott's svidence over to an independent committee for its opinion, FDA rejected the petition, stating that "the data submitted . . . do not establish that cyclamate acid, calcium cyclamate, and sodium cyclamate are safe for their intended use". Undeterred, Abbott has demanded a public hearing, and will probably press its case in court if necessary. The final curtain has still not dropped on the longest-running food additive farce in Washington.

● The Presidential election campaign drones on, with scarcely any mention of scientific, environmental or related issues (the only exception being some discussion of ways to curb nuclear proliferation). But at least the candidates' thoughts on three matters of interest to the scientific community

have appeared in print.

The current issue of *Physics Today*, a magazine published by the American Physics Society, contains the responses of President Ford and Jimmy Carter to three questions put to them by the APS president, William Fowler. Fowler asked for their views on the role of science advisers in the White House, on national energy needs and the nuclear power programme, and on federal support for basic and applied science.

On the role of scientists in White House policy-making, Ford referred to the fact that Congress has recently approved legislation he introduced to re-establish a White House Office of Science and Technology Policy (OSTP). Noting that the director of the office will be the President's science adviser, Ford stated that he would "be one of my senior advisers and will also provide advice to other senior advisers". Carter offered no details of the role of his science adviser, though he argued that "the office of science adviser to the President should be ungraded immediately to provide a permanent and high-level relationship between the White House and the scientific community".

The two candidates differed most in their replies on energy policy. Ford contended that the use of both coal and nuclear power should be expanded, called the safety record of nuclear plants "outstanding", said he had increased federal spending on safety technology, and noted that his Administration was then in the middle of a review of its nuclear policy.

Carter said that no leadership now existed in the White House on energy planning. He said strong emphasis should be placed on energy conservation, arguing that "50% of our energy is wasted". On nuclear power, Carter stated that "we must make every effort to minimise our dependence on nuclear energy", though he emphasised that he does not support a nuclear moratorium. He said he would shift research priorities from nuclear power to conservation and non-nuclear options.

On federal research support, Ford said he had increased budgets for basic research and criticised Congress for reducing his budget proposals for the National Science Foundation. Carter noted that the federal government "has a crucial role to play in supporting development of new technologies which address national priorities".

Colin Norman

USSR

● The aborted Soyuz 23 mission, although failing to accomplish any of its scientific objectives, has nevertheless won itself a place in the history of Soviet space flight by being the first major mission publicly acknowledged to have failed.

Hitherto, there has been a consistent policy for disguising failure, malfunction of lunar or interplanetary probes being disguised as part of the cover-all Kosmos programme (Kosmos 96 and 111, for example, were almost certainly originally intended as Venus and lunar probes respectively); a manned mission which returned to earth prematurely was explained in the media as having "successfully carried out its objectives". If, as in the case of the nightlanding of Soyuz 15, early return involved some special hazard, it was then claimed that the testing of recovery techniques under these difficult conditions was itself part of the programme. Even the tragic conclusion of the first Salyut–Soyuz mission, when the returning cosmonauts perished on re-entry, was offset by official pronouncements that all experiments had been successfully concluded and that scientifically speaking the mission had been entirely successful.

This reticence regarding setbacks was not lifted to any great extent in the preparatory work for the joint Apollo–Soyuz flight, for which the Soviet planners supplied information on a "need-to-know" basis, and there was relatively little discussion of past missions. Exchange of scientific information is a major concept of detente; the effective acknowledgement of difficulties both in the mission of Soyuz 23 to the Salyut 5 spacelab and in its return and landing may prove to be a sign of greater openness on the part of the Soviet space planners.

● The Far Eastern Science Centre of the Soviet Academy of Sciences is failing to make a proper practical contribution to the national economy. That, at least, is the finding of a special enquiry by the People's Monitoring Committee of the USSR. Over the past five years the Institutes of the Centre have carried out research on "more than 400 topics of a theoretical and practical character", notably in geology, vulcanology and marine resources. But, says the report, there were serious "derelictions" in the institutes' work which have prevented the proper utilisation of scientific potential and resources.

Too little is being done to apply the results of completed research to the national economy; during the last Five Year Plan, only 20 projects "with a confirmed economic effect" were realised, and out of 142 proposed inventions, only 42 received patents. The majority of research resulted merely in publications and "unplanned-for" monographs. The Centre is therefore deemed to be failing to carry out one of its fundamental functions, that of coordinat-

ing research "within the framework of the regions".

The development of the local resources and potentialities of Siberia and the Soviet Far East was the main reason for the founding of local Filials of the All-Union Academy, and the combining of existing Institutes into "Science Centres". Thus the Buryat Filial of the Siberian Branch of the Academy of Sciences of the USSR, which has recently celebrated its tenth anniversary, concentrates on solid state physics, rare-metal chemistry, and polymer synthesis—all in the name of local development—as well as the "propagation of electromagnetic waves in the conditions of the Trans-Baikal area", which, it is claimed, has brought television to more than 80% of the inhabitants of Buryatia.

The criticism of the Far Eastern Science Centre—in contrast to a practically-oriented body such as the Buryat Filial—is fully in accordance with the directives of the current Five Year Plan, which, like its predecessor, stresses the concept that research should serve the economy. In many cases, however, this demand is little more than a slogan, invoked in the reportage of all major projects, even when, as in the case of an interplanetary probe, the practical benefit to the economy seems somewhat tenuous.

As Mr Brezhnev reiterated in justification of basic research at the 25th Party Congress, "There is nothing more practical than a good theory". The findings of the investigation on the Far Eastern Science Centre suggest, however, that, at least in the lower echelons of the scientific establishment, a greater emphasis on immediate practical applications may from now on be obligatory.

● Construction of the first full-scale geothermal power station, at the foot of a large volcano on the Kamchatka peninsula, is about to begin. Kamchatka and the Kuriles form the only region of active vulcanism in the USSR: during the past year eight new volcanoes have appeared on the Kamchatka peninsula, one of which is continually active, while on Matua Island in the Kuriles an eruption of the Sarycheva volcano began on September 23, emitting a column of lava bombs, ash and steam, with explosions every $1\frac{1}{2}$–2 minutes and an ash cloud so dense that vulcanologists could not approach the island for eight days.

The possibility of using geothermal waters has long been considered a promising energy source for the development of local mineralogical and fish-processing industries. A pilot generator, at Pauzhetka, has been in operation for some time. According to information presented at a recent All-Union Symposium held in Kamchatka and the Kuriles on the utilisation of geothermal heat-sources, "There are quite sufficient high-temperature resources for geothermal stations with a total capacity of 300 MW".

● The pilot U-25 MHD-generator, which came into service in 1971, and which still remains the only one of its kind in the world, is to receive a new superconducting magnetic system, developed at the Argonne National Laboratory in the USA. In exchange, American scientists will have an opportunity of carrying out research on the U-25 jointly with their Soviet colleagues. MHD power generation is one of the main fields of Soviet-US cooperation, and the supply of the new magnetic system will form a major step in the implementation of intergovernmental agreements. The Director of the High-Temperature Institute of the Soviet Academy of Sciences, Academician A. Sheindlin, says results from research on this generator are of particular significance because it has already been in operation for some 4,000 hours, for 200 of which it has been supplying electric current direct to the Moscow grid.

Vera Rich

CANADA————————————————

Rethinking nuclear policy

A contribution to Canada's nuclear debate has come from the head of the Science Council of Canada. David Spurgeon reports from Ottawa

THE executive director of the Science Council of Canada, John J. Shepherd, has called for the federal government to set up a concentrated, focused, nuclear industrial strategy. The call, his most recent public pronouncement, indicates how much he has done to shape a new, bolder, more public and more independent role for the Council, something it has sought over this past concluding year to its first decade.

Shepherd came to the council from industry, where he was chairman of a successful high technology instrument firm, and he brought with him the vigorous, pragmatic approach one would expect from such a background. So his public statements tend to be hard-hitting and to the point. This latest, which was contained in an article in *The Financial Post*, was no different.

Shepherd acknowledged that there are serious matters still to be resolved regarding the nuclear power issue in Canada—like other industrialised countries, Canada has become locked in debate over questions like the safety and security of reactors and the disposal of radioactive waste. But it stated that nuclear power in Canada "is a fact —and it would be foolish not to take advantage of the opportunities it offers." And it went on to point out that the size of the proposed nuclear power programme in Canada in future is very large: an estimate of 70 nuclear power units by the end of the century is decidedly conservative, which means the market will be an average of $1,500 million a year for the next 25 years.
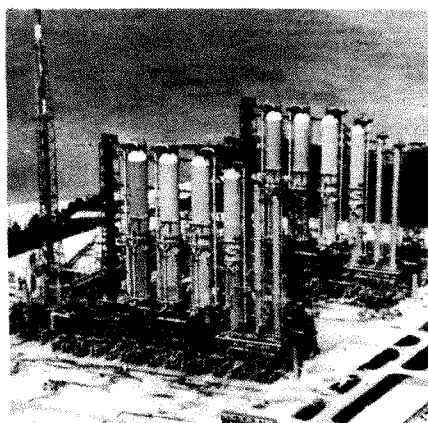
This means, said Shepherd, that "it is imperative that we devote a good deal of attention to planning." Three sectors are involved in the Canadian nuclear programme: Atomic Energy of Canada Ltd (AECL), a Crown corporation that carries out research, development and engineering; the provincial electrical utilities, which operate the plants; and industry, which does the manufacturing. "It is painfully clear," said Shepherd, "that industry has so far been unable to carry its weight in this arrangement."

This comment was nothing new: as far back as the 1960s, AECL officials were saying the same thing. But Shepherd went on to say that industry's contribution has been rendered ineffective by cancellations or postponements in nuclear plant construction, lack of a steady stream of nuclear projects, low-volume ordering and low profit margins. Those wanting to break into the market, particularly in some of the specialised instrumentation areas, were frustrated by piecemeal orders. What is required, he added, is a "mixed nuclear consortium—comprising electrical utilities, AECL, and industry", and for this to happen, several changes would have to be made.

Electrical utilities would have to alter their construction philosophies and permit others to play a greater coordinating role. AECL would have to hand over to the new consortium its Power Projects group, which carries out its engineering functions. And industry would have to accept new responsibilities. If AECL's heavy water production activities were also transferred to such a consortium, it would leave only its original research and development function. Under another name this could become the institutional focal point for a major thrust in energy research development, "for example, along the lines of the US Energy Research and Development Administration. Such an expansion of responsibilities might also contribute to the initiation of a national energy policy."

Shepherd again made it clear he thought the domestic market for nuclear power station more important to Canada than foreign ones, and referred to losses associated with a nuclear sale to Argentina. Others have pointed to difficulties Canada has had with sales to countries like South Korea, India and Pakistan. If Canadian industry could not grasp the opportunities presented by the domestic nuclear market, "it should not complain when government fills the vacuum." □

Bruce heavy water plant, Ontario

IN BRIEF————————————————

Nobel prizes announced

The $160,000 Nobel prize for medicine will be shared by Professors Baruch Blumberg and Carleton Gajdusek. Dr Blumberg is professor of Medical Genetics at the Institute for Cancer Research of the University of Pennsylvania and Dr Gajdusek works at the National Institutes of Health, Bethesda; both men did their prize-winning work at the National Institutes of Health. Both prizes are for research in virology, Professor Blumberg's for discovering Australia Antigen, a particle associated with serum hepatitis (hepatitis B), and Professor Gajdusek's for the fundamental research on kuru, the slow virus disease that was prevalent in the cannibalistic Fore tribe of New Guinea.

The prize for physics goes to Professor Burton Richter of Stanford and Samuel Ting of MIT for their work on the J/ψ particle. The particle, discovered simultaneously and independently by the two researchers in 1974, has opened up new realms of investigation with the new property of matter known as 'charm'.

The prize for chemistry is awarded to Professor William Lipscomb of Harvard for his work on boranes. The bonding of these compounds was long a puzzle according to conventional valency ideas; Lipscomb in the 1950s took the new multicentred-bonding theory, predicted borane structures and used elegant X-ray crystallography to show that some of these structures were actually cages, one even an icosahedron.

2,4,5-T production ended

Britain's only producer of 2,4,5-trichlorophenol, the chemical being manufactured at Seveso when the poison TCDD was accidentally released, has decided not to recommence production. The company, Coalite and Chemical Products Ltd, of Bolsover, Derbyshire, stopped production in August "to make 110% sure" of its safety measures. Coalite had operated with more stringent safety measures than the Italian plant, but after the Seveso accident UK Health and Safety inspectors recommended even more precautions. The company has blamed over-sensational publicity of the Seveso accident for its decision.
● A list of 721 highly poisonous substances has been prepared by the Ministry for the Environment of the West German State of North Rhine Westphalia. Included are all chemicals with an effect similar to that of TCDD, some even more poisonous and some less poisonous but still potentially lethal.

Argentinian release

Reports coming out of Argentina indicate that most of the employees of the Atomic Energy Commission who have been in prison since April (*Nature*, October 7, page 452) have now been released. Dr T. Victoria, whose brother highlighted their problems, is now in Belgium; the rest are still in Buenos Aires. There is, however, still no news at all of Antonio Misetich, a one-time MIT researcher.

UK physicists' concern

The UK high energy physics community is becoming increasingly concerned at the prospect of financial setbacks to its research effort. This became clear last week with the emergence of attempts to organise its members for concerted action. These follow the Science Research Council's recent urgent request to large laboratories for information on possible early cutbacks in expenditure.

One possibility being canvassed is that there should be a letter-writing campaign to ministers, MPs and the Advisory Board of the Research Councils, which has pursued a deliberate policy of squeezing big science. Another is that the most distinguished members of the nuclear and high-energy physics community might be able to agree on some form of corporate action to defend their interests.

The government's cash limits doctrine, which with a depreciating pound strains the SRC's international obligations, has precipitated the crisis.

FBR decision postponed

Britain's decision whether to build a demonstration commercial fast breeder reactor, due this autumn, has been delayed to give more time for public debate. Mr Anthony Wedgwood Benn, the UK Energy Secretary, has also indicated that the questions put last week to the Nuclear Installations Inspectorate concerning the fast breeder are designed to assess its margins of safety independently for the public's benefit. France's Phénix prototype fast breeder at Marcoule closed down recently for a period of weeks because of a leak in one of its heat exchangers.

Ariel V's birthday

The UK's Appleton Laboratory celebrated the Ariel V satellite's second anniversary on October 15. Ariel V, launched off the coast of Kenya, is controlled from the Appleton Laboratory which rapidly transmits data to experimental groups at the universities of London (University College and Imperial College), Leicester and Birmingham; the Goddard Space Flight Center in Maryland also operates one of the experiments.

DOES science, or do scientists, have a special responsibility to society? A conference held early in October in Florence, Italy, and organised by the Fondazione Internazionale Menarini, considered this subject from a number of angles. Experts from many countries gave papers on the problems arising from genetic engineering, ecological contamination of the biosphere, world food shortages, safer drugs for better therapy, the specific needs of developing countries and population overgrowth. The speakers identified many fields in which scientific research and its application obviously have an important part to play. They suggested that the public and their rulers often underestimate the contributions that scientists may make. But the general conclusion seemed to be that though scientists should be more vocal about their possible value, they should generally advise their rulers and should not formulate policy. In fact, they should continue to be "on tap, not on top".

Most speakers stressed the importance of work aimed at solving practical problems of food and health, and the need of more support for such investigations. However, others suggested that the important and soluble problems might not be so easy to identify, and they justified their efforts in more basic fields. For instance, it was said (predictably, by scientists with world reputations in the subject) that we know so little about the processes going on in the oceans that all manner of apparently-academic studies are justified in the hope that we may, eventually, have the knowledge to control marine pollution. It was interesting to hear views so reminiscent of the "Haldane principle" on which government-supported research in Britain was based until its recent reorganisation.

I could not help wondering how personally responsible were the scientists who were discussing responsibility. I found it interesting to observe the behaviour of many of the participants at this meeting. They had been transported freely from the ends of the earth, to be lavishly entertained by their Italian hosts, not only to the best of the food and wine of the country, but also, between sessions, to the art and music of Florence. I regret to have to report that I have seldom attended a meeting at which such a substantial number of the delegates put in such a poor attendance. Many appeared to feel that so long as they read a twenty minute paper (replete, in many cases, with material familiar to their meagre audience) and listened to a few of the other speakers at the same session, they were free to slip off to attend the "ladies" sightseeing tours or to go shopping or indulge in other non-scientific activities. This hardly seemed an example of scientific responsibility; surely delegates attending a conference at someone else's expense should be prepared to attend most of the meetings and contribute to the proceedings?

So is doing good work in the laboratory and attending meetings conscientiously the sum total of the scientists' responsibility to society? Our Marxist colleagues do not think so; it was a relief at the Florence meeting to be spared their diatribes urging the adoption of political dogmas as the acme of scientific fulfilment. I do not think I am alone in my belief that scientists are responsible not only for their own work, but for trying to ensure that this work, and that of their colleagues, is effectively organised and applied. If we are right in our views that many recent developments in scientific organisation have been harmful to both science and to society, it is our responsibility to try to have a better system adopted. We should do this no matter how unpopular it may make us with some of those in authority, and even though, under the present system, we depend on those same authorities for both support and (in the case of our younger members) the furtherance of our careers.

On responsibility

KENNETH MELLANBY

correspondence

Not a MAFF gaff

SIR,—The attack by Kenneth Mellanby in your issue of September 16 (p186) quotes recent newspaper interpretations of reports by the Ministry of Agriculture, Fisheries and Food (MAFF) but does not quote from the reports themselves, which are a strictly factual account of our food supply The MAFF evaluates Britain's food supply at two different levels The first provides an assessment of the total food available for human consumption in the UK from home production and imports, after deductions have been made for exports and non-food uses, the second (the National Food Survey) provides an assessment of food bought at the retail level by different types of household. Professor Mellanby's article appears to be based on newspaper comments on the former, which was published in *Trade and Industry* on August 27 of this year (p596)

Value judgements are not appropriate to these published accounts, and the MAFF nowhere suggests in them that "Britons eat too little food of any kind", or that "the quality of food consumed shows a steady decline", or that "the consumption of more animal protein leads to better nourishment" Rather, it is clear that an average energy value of 12 16 MJ (2,910 kcalorie) per person per day is substantially more, not less, than any recommendation from the United Nations' Food and Agriculture Organisation. Furthermore, the most recent report of the UK National Food Survey Committee shows that the food purchases of British households provide substantially more of almost every major nutrient than is recommended by the UK Department of Health and Social Security, and that the nutritional quality (estimated as amounts of nutrients per 1,000 kcalorie) is in several respects increasing.

H. C. PEREIRA
Chief Scientist,
Ministry of Agriculture, Fisheries and Food,
Whitehall Place,
London SW1A 2HH, UK

Tobacco substitutes

SIR,—Tobacco substitutes cannot be so neatly classified as the article by Allan Piper implies (*Nature,* September 2, p2); nor is it relevant, as the article might also imply, to predict the toxicological characteristics of tobacco substitutes from smoke chemistry data.

Tobacco itself, of course, has both 'organic' and 'inorganic' components Both NSM and Cytrel, which are currently the subjects of submissions to the UK Hunter Committee, are 'organic' to the extent that each contains cellulose derivatives as combustible fuels, and glycerol as a humectant; and both these tobacco supplements are also 'inorganic' to the extent that each contains a proportion of mineral fillers.

Both NSM and Cytrel produce smoke, the composition of which differs considerably from that of tobacco smoke; and to rely upon the chemical composition of smoke to predict toxicological characteristics, as Piper's article can imply, is unwise. The fact is, and there is ample evidence for this from many other aspects of toxicology, that our understanding of the biological effects of chemical compounds, acting alone or in combination. is quite insufficient to allow any such prediction to be made Instead, reliance must be placed primarily upon a comparison with tobacco in terms of tests which provide direct measurements of biological activity relevant to diseases associated with cigarette smoking, and upon other toxicological tests to afford an acceptable assurance in respect of risks to health not associated with cigarette smoking. This need is very clearly recognised in the first report of the Independent Scientific Committee on Smoking and Health. The results of such direct comparisons using NSM have been most reassuring.

T. R. C. REYNOLDS
New Smoking Materials Limited,
Wilmslow, Cheshire SK9 1QB, UK

Limiting parameters

SIR,—During the past decade or more, articles published in scientific journals have been increasingly disfigured by the use of the word 'parameters' to include any variables that can be measured (*Nature,* September 16, p387) This practice. though discouraged by one learned society (Medical Research Society), now seems to be blessed by the Shorter Oxford English Dictionary (OED)—(see Addenda p2649). The added meaning is "a distinguishing or defining characteristic or feature esp. one that may be measured or quantified, 1962" that is, a variable.

The 1964 edition of the Shorter OED gives several meanings of the word 'parameter' specific to certain disciplines such as mathematics, astronomy and crystallography, but defines its general meaning as "a quantity which is constant (as distinct from the ordinary variables) in a particular case considered, but which varies in different cases. esp. a constant occurring in the equation of a curve or surface, by the variation of which the equation is made to represent a family of such curves or surfaces, 1852" This definition seems to me (with relics of schoolboy Greek to conform with παραμετρεω, "to measure one thing by another, to compare, Plato" (*An Intermediate Green–English Lexicon,* Liddell and Scott, 1955).

Oxford lexicographers may exculpate themselves with the defence that they merely record current usage in both spoken and written form (even though it tends to perpetuate a nonsense that a parameter may be sometimes a constant, at others a variable). But do editors of scientific publications, which require exactitude for permanent records, need to be so fickle? If the Editor of *Nature* were to insist that in his journal the word 'parameter' in the general sense shall have the 1852 meaning, would not editors of all journals less widely read soon follow suit, to the great benefit of all discerning readers?

This clarification of the current fog would not resolve the underlying difficulty that there is a persistent demand for a word to embrace the measurable and estimable quantities of a system, for which the word 'parameters' is now fashionably used If a factor is measurable, there is little harm in referring to it as 'a measurable'; but if it is a derived number like the constants, a or b (parameters) of the equation of direct proportionality of the two variables x and y ($y=a+bx$) it is not directly but only indirectly 'a measurable' In the collective sense measurables plus estimables are 'quantifiables' or 'quantities', or 'numerables'.

J. F. LOUTIT
Medical Research Council,
Harwell, Didcot,
Oxford, UK

news and views

Molecular heterogeneity of the β⁰ thalassaemias

from Edward J. Benz Jr

THE β thalassaemias are characterised by an inherited reduction (β^+ thalassaemia) or total absence (β° thalassaemia) of β-globin chain synthesis. There is general agreement that defects in β-chain synthesis occurring in β^+ thalassaemia are due to corresponding reductions in the amount of RNA coding for β chains (Nienhuis and Anderson, *J. clin. Invest.*, **50**, 2459; 1971; Benz and Forget, *J. clin. Invest.*, **50**, 2755; 1971; Housman *et al.*, *Proc. natn. Acad. Sci. USA*, **70**, 1809; 1973; Kacian *et al*, *Proc. natn. Acad. Sci. USA*, **70**, 1886; 1973). But as indicated by the recent paper of Ramirez *et al.* (*Nature*, **263**, 471; 1976) the β° thalassaemias may be far more complex and heterogeneous with respect to their precise molecular origins.

Several genetically and clinically distinct haemoglobinopathies are associated with the absence of β-chain synthesis. The "typical" or "high A_2" type of β° thalassaemia represents the extreme form of β^+ thalassaemia; β-chain synthesis is totally absent and haemoglobin A_2 ($\alpha_2\delta_2$) levels are elevated. In both $\delta\beta$ thalassaemia and hereditary persistence of foetal haemoglobin (HPHF) δ and β-chain synthesis are totally absent. In HPHF the balance between alpha-chain and non-alpha-chain synthesis is preserved by efficient synthesis of the γ chains of haemoglobin F ($\alpha_2\gamma_2$). In $\delta\beta$ thalassaemia, γ-chain synthesis is insufficient to compensate for the absence of β-chain synthesis, resulting in a thalassaemia phenotype. Finally, Hb Lepore syndrome is characterised by absent δ and β-chain synthesis, incomplete compensation by γ-chain synthesis and production of small amounts of Hb Lepore ($\alpha_2[\delta\beta]_2$). The $\delta\beta$ chain is a fused chain consisting of the N-terminal regions of δ chains and the C-terminal regions of β chains. Hb Lepore is thought to arise by an unequal crossover event during meiosis, resulting in deletion of intact δ and β-chain genes and creation of a fused $\delta\beta$ locus.

Conconi *et al.* (*Nature*, **238**, 83; 1972) have suggested that the high A_2 form of β° thalassaemia is hetero-

geneous at the molecular level. In their studies, reticulocyte polyribosomes from patients with β° thalassaemia from the Ferrara Valley region of northern Italy synthesised no β-globin chains when incubated with postribosomal supernatant fractions prepared from the same Ferrara-type β°-thalassaemia reticulocytes. But "induction" of some β-chain synthesis was reported when these polyribosomes were incubated with non-thalassaemic supernatant. Conconi *et al.* (*Nature*, **254**, 256; 1975) later reported data suggesting that intact reticulocytes from these patients synthesised some β globin after therapeutic blood transfusion. In contrast, non-thalassaemic postribosomal supernatant fractions did not induce β-chain synthesis by polyribosomes obtained from non-Ferrara β-thalassaemic reticulocytes (Rowley and Kosciolek, *Nature new Biol.*, **239**, 234; 1972). Conconi *et al.* proposed that the β mRNA in Ferrara thalassaemia was normal; the primary genetic defect was thought to be a deficiency of some factor essential for β-mRNA translation.

Direct studies of globin mRNA translation in heterologous cell-free protein-synthesising systems by Benz *et al.* (*Blood*, **45**, 1; 1975), Kan *et al.* (*Proc. natn. Acad. Sci. USA*, **72**, 5140; 1975), and Tolstoshev *et al.* (*Nature*, **259**, 95; 1976) have subsequently demonstrated absence of translatable β-chain mRNA in non-Ferrara forms of β° thalassaemia. Ramirez *et al.* have now translated globin mRNA isolated from Ferrara patients before, and 22 d after, blood transfusion. No synthesis of β globin was observed when these mRNA fractions were incubated in a cell-free system fully supplemented with all the factors necessary for efficient β-chain synthesis. All forms of β° thalassaemia studied directly thus seem to be characterised by absence of functional β mRNA.

Total absence of functional mRNA could result from absence of the β-chain globin gene (gene deletion), failure of the β-globin chain to function as a template for mRNA syn-

thesis (transcription), or synthesis of structurally abnormal mRNA incapable of functioning in protein synthesis. Evidence favouring each of these mechanisms has now been reported. The paper of Ramirez *et al.* typifies the diversity of results obtained. Each laboratory studying this problem uses the same general methodological approach. Purified synthetic DNA copies (cDNAs) complementary to human α and β mRNA are isolated and used as molecular probes in cDNA–mRNA or cDNA–DNA hybridisation assays designed to quantify the relative amounts of α and β mRNA or DNA present in the sample being studied. Ramirez *et al.* have demonstrated that the β gene is present in high A_2 type β° thalassaemia but absent in HPHF, confirming recent work by others (Tolstoshev *et al.*, *op. cit.*; Ottolenghi, *et al.*, *Proc. natn. Acad. Sci. USA*, **72**, 2294; 1975; Kan *et al.*, *Nature*, **258**, 162; 1975; Forget *et al.*, *Cell*, **7**, 23; 1976). Ramirez *et al.* also report the important new finding that the β-chain gene is absent from the genome of patients with homozygous $\delta\beta$ thalassaemia. The authors cite unpublished results by Ottolenghi *et al.* (submitted for publication) from which the same conclusions were reached. There is thus general agreement that at least two genetic mechanisms account for different haemoglobinopathies associated with absent β-chain synthesis gene deletion in $\delta\beta$ thalassaemia and HPHF, and the presence of a gene which does not code for active β mRNA in high A_2 type β° thalassaemia.

There is considerably less agreement about the lesion in β-mRNA metabolism responsible for absent β-chain synthesis in patients with high A_2 type β° thalassaemia. Forget *et al.* (*Nature*, **247**, 379; 1974) and Tolstoshev *et al.* (*op. cit*) have reported virtual absence of β-chain mRNA in non-Ferrara patients with β° thalassaemia. In contrast, Kan *et al.* (*op. cit.*) studied two Chinese patients with β° thalassaemia and found nearly normal amounts of apparently intact β-chain mRNA. Ramirez *et al.* now report that reti-

culocyte RNA from Sicilian patients with β^0 thalassaemia contained significant amounts of β-mRNA which hybridised efficiently to the β-cDNA probe. The hybridisation behaviour of mRNA from patients with Ferrara-type thalassaemia was much more complex. Some "β-like" mRNA was detected, but this mRNA could protect only 40-50% of the β cDNA from digestion by the single stranded specific S_1 nuclease used in the hybridisation assay. The β-like mRNA being detected was either deficient in major portions of the sequence represented by β cDNA, or possessed a base sequence so grossly altered that very poorly matched hybrids were formed.

In summary, published data suggest that β^0 thalassaemia and related haemoglobinopathies can arise from lesions in virtually every step of globin mRNA metabolism: gene deletion (in HPHF and in β^0 thalassaemia), presence of the gene but total absence of β mRNA, presence of grossly abnormal β mRNA in Ferrara-type thalassaemia, and presence of mRNA with normal hybridisation properties but no template activity. Unfortunately, it is not clear at present whether the diversity of results reported reflects the true heterogeneity of these syndromes or, to some extent, methodological differences among the various laboratories investigating this problem. For example, Forget *et al*, and Kan *et al*. obtained very different results when they analysed the same mRNA sample from a Sicilian patient with β^0 thalassaemia. Forget *et al*. reported near absence of β mRNA while Kan *et al*. found 30–70% as much β mRNA as α mRNA. Therefore at least some of the differences reported may be due to technical factors rather than to any fundamental heterogeneity among the β^0 thalassaemias.

Many theoretical and practical considerations demand cautious interpretation of the data obtained from different laboratories. Reticulocyte counts tend to be rather low in β^0 thalassaemia and blood samples must often be transported from endemic areas to the laboratories equipped for hybridisation analysis. One is thus dealing with very small amounts of RNA, isolated, in many cases, from blood subjected to variable conditions of storage and transport. More importantly, the reticulocyte represents only the end stage of erythropoiesis; the mRNA levels measured in reticulocytes may not reflect accurately the true status of mRNA metabolism during erythropoiesis. For example, if all patients with a high A_2 type β^0 thalassaemia produced an unstable β-mRNA molecule, the different results obtained in different laboratories might merely reflect the

"age" of the reticulocytes isolated at a given time from a given patient, contamination of reticulocytes by circulating nucleated red blood cells, variations from patient to patient in the degree of mRNA instability, or differential losses of α and β mRNA during storage of the sample. Since "β mRNA" content is in fact measured as the ratio of β/α-mRNA sequences, these factors could produce either over- or under estimates of β-mRNA content, depending on which message (α or β) was more severely affected.

Close homology exists among the amino acid sequences of β,γ (39/146 differences from β), and δ (10/146 differences) globin chains. Partial cross-reaction among the mRNA and cDNA sequences specific for these chains is known to occur unless incubation conditions are carefully designed to eliminate cross-hybridisation. In all the aforementioned studies the hybridisation conditions were designed to eliminate γ–β cross-hybridisation. It is, however, much more difficult to deal with the potential problem of δ–β cross-hybridisation. It has not yet been possible to isolate δ mRNA, estimate the "normal" amounts present, or determine whether δ–β cross-hybridisation occurs in "stringent" hybridisation conditions. Finally, some laboratories use cDNA probes which are incomplete copies of their mRNA templates or are significantly contaminated (10–30%) with other globin-specific cDNA sequences. Although such probes are useful for detecting the presence or absence of β mRNA, they are less reliable for quantifying the precise amounts present.

Although the above technical reservations are significant, there is little doubt that some molecular heterogeneity exists among different patients with β^0 thalassaemia. For example, the β mRNA detected by Kan *et al*. in reticulocytes from Chinese patients had thermal denaturation profiles (melting temperature curves) indistinguishable from those of normal β mRNA. These patients, at least, seem to have some true β mRNA. Similarly, Benz, Forget, and coworkers (*Clin. Res.*, **23**, 269a; 1975, and manuscript in preparation) have analysed mRNA from a number of patients with β^0 thalassaemia using highly purified full length cDNA probes. In some of these patients the absence of message could be demonstrated, while in others the presence of at least small amounts of β-like mRNA was observed. In these studies, a patient with Hb Lepore syndrome had mRNA which hybridised to the β probe, suggesting indirectly that the δ–β mRNA (and, presumably, δ mRNA) can cross-hybridise with β cDNA even in "stringent" conditions.

Precise definition of the genetic lesions in different patients with β^0 thalassaemia is important because each of these defects represents a potentially important clue for elucidation of the poorly understood pathways of mRNA synthesis, processing, transport and stability. Further progress will require detailed investigation of mRNA metabolism in nucleated erythroblast precursors. It is to be hoped that there will be free exchange of samples so that general agreement may be reached as to the final results obtained. It is crucial that the results obtained reflect only differences in the state of the β^0 thalassaemia genes rather than differences in the state of the art in various laboratories.

The thalassaemia syndromes have been unique among human disease states because of the facility with which the techniques and concepts of molecular biology can be applied to explain the genetic and biochemical origins of the clinical disorder. Further study of the β thalassaemias should continue to yield valuable information about the factors controlling the metabolism and stability of globin mRNA in normal erythroid cells. ☐

Changing climate of the North Atlantic

from P. M. Kelly

THE debate over the causes of climatic change is slowly shifting from lively, but somewhat unproductive, discussion of the theoretical likelihood that any climatic control, such as solar activity or ocean–atmosphere interaction, could affect climate, to discussion concerning the relative importance of selected controls that have been convincingly shown to have affected climate. This synthesis has been hampered by the reluctance of some investigators to examine the assumptions upon which their chosen method of analysis is based and to consider the extent to which their conclusions are determined by such assumptions. For example, the results of some numerical modelling experiments tend to stress the importance of internal fluctuations in the ocean–atmosphere system and to minimise the importance of external factors such as solar activity, simply because these external factors are assumed constant. Similar considerations apply to all investigators who concentrate on observational evidence related to one external factor alone. The causes of climatic change are complex and it is unlikely that any one approach to the problem, or any one climatic control, will provide a full answer. To assess

WHY don't platelets stick to the surface of healthy blood vessels whereas they will stick to foreign surfaces such as glass or to damaged or diseased blood vessels? The recent discoveries of two naturally occurring compounds have revealed much about the underlying chemical mechanisms and may provide an answer to this intriguing problem.

In 1975 Samuelsson and his group at the Karolinska Institute in Stockholm discovered that when platelets stick to each other (or aggregate) they release a substance called thromboxane A₂ (*Proc. natn. Acad. Sci. U.S.A.*, 72, 2994). This highly unstable compound (half-life in water about 30 s) is formed from the same precursor as the prostaglandins (arachidonic acid through an unstable endoperoxide, prostaglandin G₂), but differs sufficiently in structure from the prostaglandins to justify the use of new nomenclature. Thromboxane A₂ (which is almost certainly identical to the rabbit aorta-contracting substance previously discovered by Piper and Vane) is a powerful stimulant of platelet aggregation. It is probable that the formation of this substance is responsible for the clumping of platelets which, if allowed to continue inside a blood vessel, may result in thrombus formation. If the blood vessel becomes completely occluded by such a thrombus, the consequent tissue damage, for example in the heart or brain, may prove rapidly fatal.

Oates and his colleagues at Vanderbilt University have reported (*Science*, 193, 1135; 1976) that thromboxane A₂ released from human platelets can constrict coronary blood vessels of the

Molecular insight into thrombosis

from E. W. Horton

pig. They postulate that platelet aggregation in areas of damaged endothelium can release thromboxane A₂ and thus cause constriction of large coronary arteries.

A more recent development is reported in this issue of *Nature* (page 663). It comes from Vane's group at the Wellcome Research Laboratories, Beckenham. In the microsomal fraction of homogenised blood vessels they have discovered an enzyme which converts the endoperoxide intermediary, prostaglandin G₂, not to the classical prostaglandins (E₂, F₂ₐ and D₂) nor to thromboxane A₂ (as occurs in platelets) but to a new substance (PGX) which has not yet been identified. PGX has two important pharmacological properties both opposite to those of thromboxane A₂. It is the most powerful inhibitor of platelet aggregation yet discovered and it dilates blood vessels. Thus the endothelial lining of blood vessels possesses its own mechanism for generating a substance which will prevent platelets from adhering to its surface or to each other. Moreover, the substance will tend to maintain the flow of blood by dilating the vessel itself.

The following sequence of events can therefore be envisaged. When platelets come into contact with endothelium of the normal blood vessel wall they release a quantity of the prostaglandin endoperoxide, which is then used as a substrate by the blood vessel enzyme to form PGX. By its inhibitory properties this substance will, in turn, counteract any tendency of platelets to aggregate, or of the blood vessels to constrict. If, however, platelets come into contact with blood vessels whose endothelium has been disrupted or destroyed by pathological changes, there will be no enzyme for the formation of PGX and so the aggregating action of thromboxane A₂ formed by the platelets from prostaglandin endoperoxide, will go unopposed leading to thrombosis and local vasoconstriction.

Since PGX, like thromboxane A₂, is an unstable compound, its usefulness as a drug to treat or prevent thrombosis in man may be limited. Nonetheless, the elucidation of its structure is of paramount importance. Once the structure is known, analogues can be synthesised with the object of finding a compound which combines the pharmacological (PGX-like) activity with greater stability. Such a drug could be used in both the treatment and prevention of thrombosis.

Thrombosis is one of the most common causes of death in the western world. Young and middle-aged men, apparently fit and in the prime of life, are frequent victims. Few really effective remedies or preventative measures are available. Until now, in spite of much outstanding research, the sequence of events which leads to thrombosis at what is popularly called the 'molecular level' has been incompletely understood. The potential value to medicine of the discovery of PGX is therefore incalculable.

---

the relative importance of the climatic controls, which is an essential step towards climatic prediction, collaboration between people using often widely divergent methods to study climatic change is highly desirable.

The recent article by Colebrook (*Nature*, 263, 576; 1976) highlights one area where such collaboration should prove fruitful. Colebrook analyses long series of oceanographic and climatic data in order to define changes in the North Atlantic current systems and ocean temperature during the past 100 yr and to suggest ways in which shifts in the ocean and atmospheric circulation are related. He concludes that, on time scales of decades, sea surface temperature changes are largely due to variations in advection associated with the changing intensity of the Gulf Stream and North Atlantic Drift currents. He states that these currents are responding to variations in the atmospheric circulation over the North

Atlantic, although the atmospheric circulation index chosen (tropical cyclone frequency) is remarkably indirect.

Colebrook also shows, contrary to other widely publicised opinions, that recent changes in the oceanic climate are not unusual in the context of the past 100 yr. In addition, he presents further evidence of the approximately 10-yr periodicity in sea surface temperature in the North-East Atlantic that has been mentioned by other investigators. He links this periodicity to changes in atmosphere pressure associated with the sunspot cycle, as given by Parker (*Met. Mag*, 105, 33; 1976). It is interesting that the connection between variations in the strength of the East Atlantic meridional atmospheric flow and solar activity is also apparent in the 24-h change in atmospheric pressure distribution following a solar flare (see, for example, Schuurmans, *Meded Verh. K. ned. met. Inst.*, 92; 1969). The effects

on climate of solar variation are gradually being accepted as genuine but of secondary importance, tending to modify the nature of the meridional rather than the dominant zonal atmospheric air flow.

Increasing evidence of frequency-dependent relationships between climatic variables is coming to light. Sea surface temperature to the west of the United Kingdom seems to be controlled by the strength of the broad Atlantic westerly air flow on time scales of decades and longer. Yet, on the shorter time scale of the 10-yr periodicity, local changes in the meridional component of the atmospheric circulation are more important. This association between sea temperature and southerly air flow in the East Atlantic on short time scales has been demonstrated by Dickson (*Dt hydrogr. Z.*, 24 (3), 97; 1971).

Investigations such as this are hampered by the lack of long climatic data series. As usual when limited data are

available, hypotheses multiply, each able to explain the observations. The contribution of such empirical analyses to the search for the causes of climatic changes can be maximised if they are used in conjunction with the theoretical approach to the same problem: numerical modelling. The numerical modelling of climatic change and the general circulation of the atmosphere is at an early stage of development; such experiments are primarily designed to test the behaviour of the models, using unrealistically large anomalies in the boundary conditions to produce an unequivocal response in the atmosphere. Until reliable coupled ocean–atmosphere m o d e l s become available, numerical models can only provide a static picture of a climatic state and cannot portray climatic evolution. Their potential, however, is great and slowly being realised. Hypotheses concerning the causes of climatic change based on empirical evidence can be tested by numerical simulation and *vice versa*.

To assess the relative importance of the climatic controls, it is necessary to specify realistic anomalies in the boundary conditions of the models. Otherwise the value of the experiment is lessened. It may indicate that a particular control can be effective, but not how effective. The use of information concerning actual changes in the boundary conditions of the atmosphere, such as contained in the article by Colebrook, should ensure that the results of numerical modelling experiments do not diverge unnecessarily from reality.

Increased collaboration between people studying climatic change within a theoretical framework and those concentrating on the observational data is highly desirable, and perhaps essential if climatology is to provide some indication of future c l i m a t i c development. ☐

Chinese mitten crab reappears in Britain

from R. W. Ingle and M. J. Andrews

THE unintentional transportation of animals by man beyond the range of their normal habitats has always interested zoogeographers. Five species of true crabs have arrived in British waters in this manner. One of these, the Chinese mitten crab, is of particular interest as its spread through Europe since its introduction from China has been well documented and because it has now reappeared in Britain after an
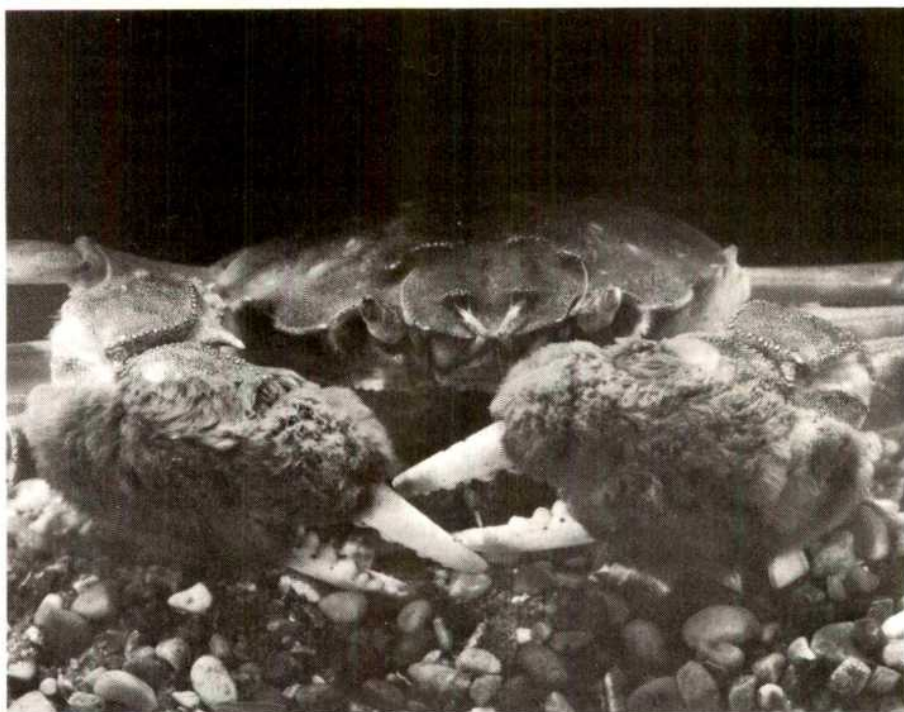
Front view of the male mitten crab from West Thurrock generating station. This species is easily recognised by the dense fur on the claws. The body colour varies from greenish grey to dark brown.

apparent absence of some 27 yr.

The mitten crab, *Eriocheir sinensis* A. Milne Edwards, is a native of temperate eastern Asia and occurs from the Fukien province of China to the western coastal region of Korea. In Korea it shows preference for rice fields near the coast and inland occurs only in rivers. Life in this temperate environment has enabled the crab to adapt to similar climatic conditions in northern Europe from where it was first reported in 1912 from the Weser River near Retham, Germany and later from other parts of the country. In the 1920s the crab spread rapidly westward and was recorded from rivers in the Netherlands, Belgium, France, Denmark and Finland. Its occurrence in Britain was first reported in 1935 when a male was found in the River Thames at Chelsea, then in 1949 a second specimen was discovered in Southfields reservoir near Castleford, Yorkshire. There have been no other reports of this crab in Britain until this year when three specimens were taken from the cooling water intake screens of West Thurrock electricity generating station in February, May and June. The West Thurrock station is situated in a brackish region of the River Thames estuary 36 km below London Bridge where the average half-tide salinity for the first half of 1976 was approximately 20 g per 1,000 g.

The mitten crab probably reached Europe in the ballasting tanks of commercial vessels, having entered the

tanks in Chinese waters as free swimming larvae and developed *en route* to be liberated as juveniles when the tanks were emptied in Germany. Studies in Europe have shown that adult crabs prefer freshwater and live in rivers but migrate into estuaries to breed. The eggs are hatched in brackish water and developing larvae move gradually back into freshwater.

This recent reappearance of the mitten crab in the Thames is interesting because all the specimens were found within a period of 5 months and in an area that has been surveyed regularly since the 1960s, during which time the crab was never reported. It is probable that these individual crabs were transported to the West Thurrock region by ships arriving from European ports (more than 100 ships from Antwerp and Rotterdam have visited West Thurrock in the past 2 yr), and does not constitute a serious invasion by the species. In large numbers mitten crabs can cause considerable damage to estuarine river banks resulting from the subsidence of their numerous small burrows in the mud. The crab is also reported to cause damage to fishing nets by cutting the mesh in which they often become entangled. There seems no obvious reason why *Eriocheir* should not become established in the British Isles and the arrival of females carrying eggs would be a significant step towards this invasion; further occurrences of the mitten crab are awaited with interest. ☐

Rheology at Gothenburg

from D. R. Oliver

The Seventh International Congress in Rheology was held at Chalmers University of Technology, Gothenburg, Sweden on August 23–27, 1976, organised by Professor Josef Kobat and colleagues (Chalmers University).

RHEOLOGY is the study of the deformation and flow of materials, but a visit to Gothenburg provided opportunities to observe the deformation and flow of travellers cheques! The subject subdivisions chosen by the organisers were polymer solids, polymer melts, fluids, polymer solutions, theory, suspensions, biorheology, methods, metals, concrete and food. This order represents a ranking in terms of numbers of papers presented: new subjects such as biorheology (13 papers) and food (4) seem to be making little impact on the combined polymer topics (149). The apparent fall in the number of theoretical papers (23) is illusory, since papers under other headings contain extensive theoretical work and one new concept in particular seems to be gaining momentum, the search for rheological equations of state consistent with thermodynamics (J. G. Oldroyd (University of Liverpool) and G. Astarita (University of Naples)). Another link between rheology and physics was provided in a historical lecture by H. Markovitz (Carnegie–Mellon University, Pittsburgh) who pointed out that 100 yr ago Boltzmann formulated his principle of superposition, according to which stresses present at a given time depend not only on the deformation at the time but also on previous deformations, the influence of which decreases with increasing elapsed time. The suggestion was made that Boltzmann's ideas should be given status comparable with those of Maxwell in the early development of theories of viscoelastic behaviour.

Flow visualisation is widely used in rheology: some unusual and beautiful extrudate flow forms were shown by N. Bergam (CIIR, Oslo) and die-entry flows, with customary flair, by Professor Giesekus (University of Dortmund). The quadruple vortex formed in a Taylor 4-roll mill was demonstrated by J. M. Broadbent (University of Wales, Aberystwyth) and the flow pattern around a sphere moving in a viscoelastic liquid by D. Sigli and M. Coutanceau (University of Poitiers). The pattern of a recoiling flow upstream of

an orifice upon sudden cessation of flow was shown by R. J. Gordon (University of Florida) and a remarkable tri-modal Weissenberg effect induced by torsional oscillation of a rod in a viscoelastic liquid by D. D. Joseph (University of Minnesota). The liquid climbing the rod builds up into three (or even four) distinct lobes before eventually collapsing.

A computational method finding increasing acceptance is the finite element system, in which the uniform grid of a finite difference network is replaced by a set of irregular elements which concentrate attention around areas of particular interest. R. Tanner (University of Sydney) obtained valuable information regarding velocity profiles at the exit end of a capillary tube and was able to predict jet swell for Newtonian fluids; N. Davids (Pennsylvania State University) and M. L. Wenner (General Motors, Michigan) used finite elements for the study of uniaxial waves in a viscoelastic rod; D. S. Malkus (National Bureau of Standards, Washington) calculated hole pressure error, and G. Geymonat (CNRS, Marseilles) and M. Raous (Poletecnico, Turin) applied the technique to the near-solid problems of the behaviour of turbine blades and to time-dependent effects in ageing concrete. The finite element method will become widely used in rheology and other branches of fluid flow.

The flow of blood was the dominant theme of the biorheology section. New viscosity concentration relationships were obtained by D. Quemada (University of Paris) and hysteresis effects in whole human blood in Couette flow studied by A. Apelblat *et al.* (Paris). These experiments showed that the aggregation of red cells is both a time- and shear-dependent phenomenon producing hysteresis loops in rotary instruments. The possible migration of red cells away from the walls should not be ignored; these effects occur strongly in flow down fine capillaries and produce the Fahraeus–Lindquist effect, that is, lower than normal apparent blood viscosity. M. Singh (IST, Madras) and N. A. Coulter (University of North Carolina) studied the latter behaviour in oscillatory flow down capillaries and reached the interesting conclusion that the effect begins at larger tube radii than for steady flow—this may help the human body considerably! In a significant paper L. E. Gelin (University of Gothenburg) showed how the rheological changes in blood following shock or trauma influence its distribution in fine capillaries. A Perspex model allowed blood to flow along capillaries and then to choose a "straight on" or "branching" path. For highly aggregated blood (such

as that following shock), the forward flow contained more red cells and the side flow less, particularly for very narrow tubes. In critically ill patients, this may lead to the total immobilisation of red cells, visible as "red shock".

Several new concepts were introduced. B. Mena and O. Manero (National University of Mexico) showed that the flow rate of a viscoelastic liquid along a tube could be increased 10-fold by the low frequency axial oscillation of the tube, producing spectacular spurts of liquid. A paper by K. Walters (University of Wales, Aberystwyth) described his "torsional balance" variant of the rheogoniometer, which started life as a squeeze film system but has become a potential rival to jet thrust as a high-shear steady-state normal stress measuring instrument.

Extensional measurements continue. L. Nicolais (University of Naples) and colleagues showed that the extensional viscosity of polymer relations containing fine glass beads in suspension is actually reduced due to the presence of the filler; conversely D. R. Oliver and R. C. Ashton (University of Birmingham) showed that the polymers added to car lubricating oils cause large increases of extensional viscosity in both uniaxial and biaxial flows.

A novel fibre foam was described by S. Turner and F. N. Cogswell (ICI, Welwyn Garden City). This fibre/resin/air matrix is easily bonded to other materials and is light and strong. The uses suggested include shock absorbent packing, fenders and riot shields—certainly a topical development. □

A hundred years ago

Interferon therapy for hepatitis

from Arie J. Zuckerman

THE major importance of persistent infection with hepatitis B virus and the carrier state can be considered first in terms of the potential dissemination of the infection to contacts by various parenteral (direct introduction into the tissues) and non-parenteral routes; and there are, at a conservative estimate, some 110 million carriers of hepatitis B surface antigen. Second, a significant proportion of infected patients and carriers have chronic active hepatitis, cirrhosis and probably primary liver cancer in certain areas of the world. Much has been written on this subject and considerable effort is being directed at passive and active immunisation. There is, however, no really effective treatment which affects the course of chronic infection with hepatitis B virus, but somewhat unexpected and encouraging results have now been reported after systemic treatment with exogenous interferon, first by Greenberg and associates (*New Engl. J. Med.*, **295**, 517; 1976) and more recently by Desmyter and colleagues (*Lancet*, **ii**, 645; 1976).

Interferon has a well known broad spectrum antiviral action, and its lack of antigenicity in the homologous species and the absence of toxicity immediately provide a compelling appeal for its use. Indeed interferon has been investigated intensively since its discovery in 1957 by Isaacs and Lindenmann, and much effort has been devoted to its application in clinical practice for the treatment of several viral infections (Merigan *et al.* in *Antiviral Mechanisms, Perspectives in Virology*, **9** (edit. by Pollard, M.), 249; Academic, 1975). Interferon is probably not a single substance but rather it represents a class of cellular glycoproteins occurring as monomers or polymers of a basic unit with a molecular weight of 12,000–20,000. It is synthesised by all vertebrates studied in response to viral infection and other intracellular parasites, polyanions, endotoxin and certain low molecular bis basic organic compounds. It is also produced as part of the cell-mediated immune response by lymphocytes.

There is experimental evidence that the viral inhibitory effect of interferon is at the level of translation of viral proteins; other data indicate that transcription of the viral genome is inhibited. But it now seems from other evidence that the viral inhibitory effect may be exerted not by interferon itself but by a cellular protein induced by interferon, and further that there is a range of proteins that can be considered

as "interferons". Interferon also has an inhibitory effect on rapidly dividing cells including tumour cells, and its antitumour potential is under investigation. Interferon can be expected, therefore, to be more effective for prophylaxis than for the treatment of established virus diseases. However, in terms of disease process resulting from continuing viral replication and spread, interferon may be useful, for example by preventing the infection of fresh cells.

Greenberg *et al.* investigated the effect of exogenous human leukocyte interferon on persistent hepatitis B virus infection associated with chronic active hepatitis in four patients. The interferon was produced by Dr Kari Cantell by stimulation of human blood buffy coats with Sendai virus. The specific activity of the interferon was 5×10^5 units per mg protein. The four patients selected for the investigation were carriers of hepatitis B surface antigen for more than 6 months, and they had persistently abnormal liver function tests and histological changes in the liver consistent with chronic active hepatitis. In addition, three of the patients had high levels of circulating Dane particle markers including DNA polymerase, core antigen and Dane particle-associated DNA. Hepatitis B *e* antigen was present in these three patients, and the fourth patient had *e* antibody (see *Nature*, **263**, 374; 1976). The administration of interferon in a dose of between 6×10^3 and 17×10^4 units per kg per d was associated with a rapid and reproducible fall in DNA polymerase activity and the other Dane particle markers in the three patients on five separate occasions. This effect was transient when the interferon was given for up to 10 d, but seemed to be more durable when the course of treatment lasted for a month or longer. Prolonged therapy was associated with the disappearance or diminution of the *e* antigen in two of the patients. The effect of interferon on the surface antigen was more variable. Short term treatment had no effect on the titre of the surface antigen, but treatment for a month or longer was associated with a significant fall in titre in two patients.

Interferon thus seems to be exerting a suppressive effect on the production of Dane particles, the presumed infectious agent of hepatitis **B**. It is not known, however, at what stage of assembly of the Dane particle this effect occurs. Furthermore, interferon has also been shown to affect both humoral and cellular immunity, and Greenberg and associates point out that the suppression of the markers associated with hepatitis B virus may have been due to a direct antiviral

action or mediated by the immune system or a combination of both. There is also the possibility that the active principle is not interferon but some other lymphokine including transfer factor (*Nature*, **258**, 14; 1975) or other cell product which is present in the partially purified preparation collected from the human blood leukocytes. But this seems unlikely in view of the essentially similar results obtained by Desmyter *et al.* (*loc. cit.*) with interferon prepared in cultures of human diploid fibroblasts using the double-stranded RNA, poly(I) · poly(C), and superinduction. This preparation was given to two chimpanzee persistent carriers of hepatitis B surface antigen and one patient with antigen-positive chronic aggressive hepatitis. The treatment consisted of seven doses of 10^7 international units of human interferon given on alternate days for 2 weeks. Interferon had no effect on the titre of the surface antigen of the two chimpanzee carriers but striking changes were observed in the hepatitis B core system. In the first chimpanzee, 15% of the nuclei of liver cells were found positive for core antigen by immunofluorescence. After treatment for 1 week, the number of positive nuclei and the intensity of staining decreased progressively during 5 weeks, and at that time 0.5% of the nuclei were positive. Then the number of positive nuclei increased rapidly reaching 30% after 1 month. The titre of circulating core antibody also increased. The second chimpanzee had no demonstrable core antigen in the liver cells at any time, but after treatment the high titre of serum core antibody fell sharply. In the patient, changes were seen in the core system. Intense core antigen staining was found in 30% of the liver cells before treatment. Immediately after treatment, 20% of the liver cell nuclei were positive at low intensity, and 5 weeks later only 3% were positive. The titre of core antibody in the serum remained unchanged.

These preliminary results obtained in short-term investigations in a few patients are promising and potentially most important, since interferon therapy may be useful in limiting the infectivity of carriers and it may perhaps eradicate chronic infection. The experimental hepatitis B vaccines currently under development are unlikely to achieve this in the foreseeable future. However, even the most ardent proponents of interferon and lymphokine therapy will urge carefully designed studies and only cautious optimism, for much remains to be learnt and accomplished before this form of treatment can be applied successfully to the many important clinical problems so clearly outlined by Merigan and his colleagues. □

review article

RNA polymerase specificity and the control of growth

Andrew Travers*

RNA polymerase is an allosteric enzyme whose activity and specificity are controlled by interaction with a nucleotide, tRNA and proteins. In E. coli the availability of these effectors should ensure that the quality of transcription is tightly coupled both to translation and to the metabolic state of the cell.

RNA polymerase is structurally one of the most complex enzymes in the bacterial cell. This key component of the machinery for the expression of genetic information is a heteromultimer of molecular weight $\sim 5 \times 10^5$ (ref. 1). Yet, in the case of the mitochondrial RNA polymerase a single polypeptide chain of molecular weight $\sim 6 \times 10^4$ suffices for the actual process of transcription[2]. The complexity of the bacterial enzyme must thus reflect additional roles in the regulation of transcription and DNA replication. One aspect of this control is promoter selection. In this review I argue that RNA polymerase is an allosteric enzyme with a variable initiation specificity and can discriminate between different types of promoter site. The molecular basis for such discrimination is the ability of the polymerase to exist in functionally distinct forms, each of which preferentially initiates at a particular class of promoter sites. One effector, the nucleotide ppGpp, whose intracellular level is tightly coupled to both protein synthesis and the metabolic state of cell[3], alters the specificity of the polymerase by converting one form of the enzyme to another. Other effectors include fmet tRNA$_f^{met}$ and EF-TuTs, both of which are intimately associated with translation. The emergent pattern of transcriptional control is thus one in which RNA polymerase itself has a central role in the determination of the quantity of the transcript. The purpose of this control is to ensure that the production of ribosomes and mRNA is balanced and geared to available energy and to the rate of protein synthesis. On this model the mutual interaction of polymerase and specificity modulators thus constitute a mechanism for the control of growth.

RNA polymerase specificity

The extent to which the direct control of RNA polymerase activity complements the operation of highly specific transcriptional controls by DNA-binding repressors and activators during normal bacterial growth has remained poorly understood. To establish a regulatory role for the transcribing enzyme in the selection of genetic information it is necessary to demonstrate both that the purified polymerase can be directly controlled *in vitro* and that mutations resulting an an alteration in polymerase structure significantly alter the pattern of transcriptional regulation *in vivo*.

One possible mode of regulating RNA polymerase activity would be the control of promoter selection. The

*Address: MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

binding of RNA polymerase to a promoter site is the first step in the synthesis of an RNA molecule and a major determinant of the initiation frequency *in vitro*. Following this initial binding there is a change in the conformational state of the DNA in the promoter region[4,5]. Once this transition has occurred the polymerase can initiate an RNA chain very rapidly[6]. The transition between the DNA conformational states is reversible and often occurs over a narrow temperature range, the midpoint of which defines a transition temperature, t. With a given set of conditions *in vitro* each promoter will have a characteristic transition temperature for opening by RNA polymerase holoenzyme. The transition temperature is dependent on factors which directly affect the structure of the DNA or which alter the affinity of the enzyme for the promoter site. Thus a change in initiation specificity of RNA polymerase can be defined as a differential alteration in t between promoters.

Experimentally two criteria can be used to measure the relative affinity of RNA polymerase for different promoters. In this context the term affinity describes the strength of a polymerase–promoter interaction which is necessary for the formation of stable complexes without any qualification as to the exact molecular requirements for this interaction, that is as to whether or not binding to the DNA double helix in the promoter region is sufficient. One criterion is the measurement of the temperature dependence of polymerase–promoter complex formation, a decrease in the affinity of the enzyme for the promoter being reflected in an increase in t. A second criterion is the measurement of the response of a particular type of RNA synthesis to ionic strength. Increasing salt concentration reduces the binding of polymerase to the template concomitantly increasing t[7]. Thus at a constant temperature a decreased affinity of RNA polymerase for a particular promoter may be reflected in an increased salt sensitivity of transcription of the cognate RNA. Neither of these criteria is ideal since the physical structure of RNA polymerase is also to some extent dependent on both ionic strength[8] and temperature[9] and it cannot be excluded that these structural changes might in themselves alter the pattern of transcription. Nevertheless used together these criteria provide useful indicators of the relative affinities of polymerase for different promoters.

ppGpp—an effector of RNA polymerase specificity

In vivo any alteration in the initiation specificity of RNA polymerase might be expected to yield a wide ranging change in the pattern of RNA species synthesised. One such

change occurs on starvation of certain bacterial strains for a required amino acid. In this stringent response the accumulation of stable RNA species, both rRNA and tRNA, is severely restricted while, in contrast, the production of certain mRNA species, for example, φ80 mRNA[10], is relatively unaffected. The reduction in rRNA accumulation is primarily a consequence of the diminished initiation of rRNA chains. Another rapid response to amino acid starvation is the intracellular accumulation of the guanosine nucleotide, ppGpp, to millimolar levels within seconds of the onset of starvation[11]. Mutants unable to accumulate ppGpp in this manner fail to shut off rRNA synthesis and are said to show relaxed control[12]. This correlation has led to the proposal that this nucleotide is the mediator of the stringent response *in vivo*[12].

There is considerable evidence that *in vitro* ppGpp can preferentially inhibit the synthesis of rRNA relative to total transcription both in crude[13,14] and in highly purified systems[15-18]. At issue is the manner in which nucleotide effects this inhibition; whether RNA polymerase itself is the target or whether additional transcription factors are required. Early studies showed that ppGpp inhibited total RNA synthesis by purified *Escherichia coli* RNA polymerase on a variety of DNA templates[15,19]. The specificity of this inhibition can be determined by comparing the effect of ppGpp on the synthesis *in vitro* of rRNA with the synthesis of φ80 RNA, an RNA species which is not stringently controlled *in vivo*, using as template the DNA of a transducing phage, φ80 d3 rrnB+ ilv+ su+7, containing both φ80 and rRNA promoters. Such a comparison shows that ppGpp increases *t* for opening the rRNA promoter by 10–15 °C; without the nucleotide *t* at 0.075 M KCl is 25–30 °C, in the presence of ppGpp *t* is 35–40 °C (ref. 17). In contrast ppGpp has little effect on *t* for opening φ80 promoters (for example, Fig. 1). ppGpp acts by inhibiting the formation of polymerase–rRNA promoter complexes and is without significant effect on the subsequent extent of RNA synthesis once such complexes have formed. Thus ppGpp clearly alters the initiation specificity of RNA polymerase holoenzyme *in vitro* in a direction which parallels the stringent response *in vivo*.

How does ppGpp alter the pattern of initiation by RNA polymerase? The experimental evidence shows that the effector alters the properties of RNA polymerase itself. To demonstrate this the enzyme is first incubated with ppGpp in the absence of nucleoside triphosphates or templates for a given time. Then the polymerase is assayed by addition to a reaction mixture containing these components but no additional ppGpp. When the enzyme is tested immediately after mixing with ppGpp an initial inhibition of all types of RNA synthesis is observed[17]. Thereafter on

further preincubation the specificity of transcription slowly changes so that the enzyme recovers almost entirely the ability to transcribe φ80 and T2 DNA but its capacity to synthesise rRNA diminishes further, plateauing at ~20% of the control value. During this change no detectable degradation of ppGpp is observed. The half time of this switch in specificity is 30–50 s. Functional removal of ppGpp by dilution reverses the change in specificity with approximately the same kinetics. Other guanosine nucleotides, for example ppG, do not produce this effect. Thus the specificity change induced by ppGpp is reversible and is specific for the nucleotide. Although the rate of change of transcription specificity is slow by comparison with other protein isomerisation constants, the immediate inhibition of polymerase activity on mixing enzyme and nucleotide suggests that this is not a consequence of a slow simple association of the two components. Rather the data are compatible with a fast association followed by a slow structural change in the enzyme. The nature of this change is discussed below.

Molecular basis of RNA polymerase specificity

The differential response of the synthesis of RNA species to ppGpp establishes that RNA polymerase by itself can discriminate between different types of promoter and that this discrimination can be regulated *in vitro*. The promoters themselves must thus differ in recognition parameters for the enzyme.

How is this discrimination effected at the molecular level? The properties of RNA polymerase isolated from *E. coli* infected with phage T4 provide a clue. After T4 infection a selective shut off of stable RNA synthesis occurs, a process which requires an unidentified T4 gene product[20-22]. In addition RNA polymerase is structurally altered. One change, termed modification[23], involves the covalent addition of an adenosine diphosphoribose residue to each α subunit[24] while in addition three small phage-induced polypeptides are found associated with the core RNA polymerase[25-27]. When core polymerase from T4-infected cells is supplemented with σ factor from uninfected cells, a requirement for rRNA synthesis[28], the reconstituted holoenzyme has a lower affinity for rRNA promoters than holoenzyme from uninfected cells[29]. Indeed the qualitative pattern of rRNA synthesis by "T4 holoenzyme" is very similar to that of normal holoenzyme in the presence of ppGpp. The transition temperature for opening the rRNA promoter is high, ~35–40 °C, and rRNA synthesis is very salt sensitive. Further the ability of polymerase to transcribe phage DNA templates is relatively unimpaired after T4 infection. The reduction in affinity of the enzyme for rRNA promoters is
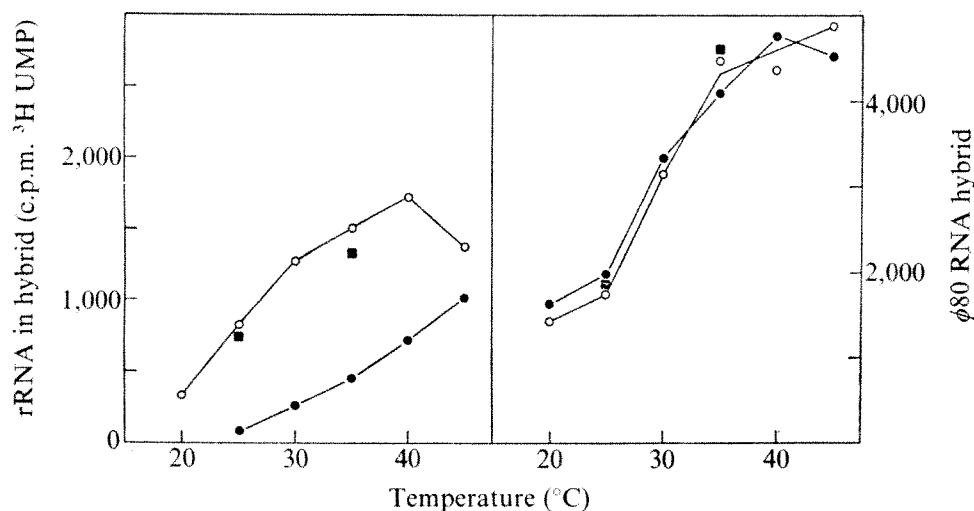
Fig. 1 Effect of temperature and 0.2 mM ppGpp on φ80 RNA and rRNA synthesised from preformed promoter–polymerase complexes on φ80 d3 DNA. Promoter polymerase complexes were formed by incubation of template and enzyme at the indicated temperature for 15′ and were assayed by the simultaneous addition of heparin and the nucleoside triphosphates: ○, control; ●, preincubation and assay in presence of 0.2 mM ppGpp; ■, 0.2 mM ppGpp added together with heparin and nucleoside triphosphates. Data are taken from ref. 17 where a full description of methods can be found.

also shown by polymerase prepared from cells infected with a T4 mutant unable to modify the α subunits[24] and is probably a function of a small phage coded polypeptide of molecular weight 15,000.

Thus at least two types of rRNA synthesis can be distinguished by purified RNA polymerases. With *E. coli* holoenzyme alone *t* for opening the rRNA promoter is 25–30 °C; that for *E. coli* holoenzyme in the presence of ppGpp and for T4 holoenzyme is 10–15 °C higher. Thus RNA polymerase can exhibit either a low or a high affinity for the rRNA promoter when that promoter is on the DNA of a transducing phage. However, when *E. coli* DNA is a template *in vitro* in conditions of limiting polymerase, rRNA synthesis by *E. coli* holoenzyme has the characteristics of rRNA synthesis by T4 holoenzyme[28,31]. The transition temperature is high and is not increased by ppGpp. This shows that *E. coli* holoenzyme can also exhibit a low affinity for rRNA promoters in the absence of ppGpp and suggests that the enzyme may exist as an isomeric mixture of two, or more, forms which differ in their affinity for rRNA promoters, and that the form for rRNA synthesis is 'used up' on *E. coli* DNA.

Theoretically RNA polymerase could exist either as a single form with a uniquely defined specificity or in a small number of conformational states, each of which would have a preferred affinity for a particular class of promoter sites. On the first model discrimination would depend solely on the relative affinity of the enzyme for particular promoters while on the second discrimination would also depend on the distribution of the enzyme between the different states.

These formal possibilities can be distinguished experimentally by template competition experiments *in vitro*. First, a system is characterised in which RNA synthesis proceeds simultaneously from two types of promoter, for example, phage and ribosomal RNA promoters. This system is then perturbed by the addition of increasing quantities of another DNA template. This DNA will cause a redistribution of polymerase molecules by competing for the enzyme. If polymerase exists as a single form with invariant specificity, the relative competition of RNA synthesised from different promotors should be independent of the competing template and depend only on the relative affinity of the enzyme for the respective promoters In contrast, if the enzyme exists in multiple forms each with differing initiation specificity the character of the competition observed should depend on the nature of the competing template. In a mixed template system containing φ80 d₁ rrnB⁺ ilv⁺ su⁺7 DNA and T2 DNA, addition of T7 DNA results in a reduction of T7 RNA synthesis and a stimulation of rRNA synthesis[32]. By contrast, addition of *E. coli* DNA reduces the synthesis of both rRNA and T2 RNA but completes rRNA synthesis to a greater degree. At the same time the residual rRNA synthesised in the presence of *E. coli* is no longer preferentially sensitive to ppGpp. Thus the character of the competition observed depends on the nature of the competing template, suggesting that RNA polymerase molecules do form a mixed population in which the enzyme can assume two, or more, recognition states. By contrast, in the presence of ppGpp the relative competition of T2 RNA and rRNA synthesis is virtually dependent of the nature of the competing template. These results also imply that *E. coli* DNA preferentially binds a form of the enzyme which has a high affinity for rRNA promoters and is selectively inhibited by ppGpp. Consequently rRNA synthesis from *E. coli* DNA by this form of the enzyme can only be detected after titration of the competing binding sites.

These arguments support the hypothesis that RNA polymerase holoenzyme exists in functionally distinct forms each of which preferentially initiates at a particular class of promoter sites (Fig. 2). Holoenzyme may exist in at least two forms, one with a high affinity for rRNA promoters, the second with a low affinity. In the presence of ppGpp and

after T4 infection the enzyme exhibits the activity of principally the low-affinity rRNA form, while retaining the ability to transcribe phage DNA efficiently. Thus the low-affinity rRNA form may have a relatively high affinity for phage promoters.

What is the mechanism of the alteration of initiation specificity by ppGpp? The nucleotide could act either by directly inactivating one class of polymerase molecules, thereby reducing the total amount of enzyme available, or by converting the high-affinity rRNA form into the low-affinity rRNA form. The kinetics of the change in specificity induced by ppGpp are compatible with conversion. Further, low concentrations of ppGpp increase transcription of phage DNA templates in conditions of DNA excess suggesting that the nucleotide can increase the amount of enzyme able to transcribe phage DNA efficiently. This type of evidence suggests that the specificity isomers of RNA polymerase are interconvertible.

The crucial question is whether the variations of initiation specificity observed *in vitro* are biologically relevant. In two instances where stable RNA synthesis is shut off *in vitro*, ppGpp accumulation and T4 infection, the affinity of RNA polymerase for rRNA promoters *in vitro* is reduced. In the latter case there is good evidence that this parallel reflects the situation *in vivo*. Recently mutants of T4 have been isolated which fail to shut off host RNA synthesis normally. RNA polymerase prepared from cells infected with these mutant phages lacks the 15,000 dalton T4-coded polypeptide (Gold, Kutter, Sirotnik and Snyder; Snustad, personal communications) which the *in vitro* studies had implied was necessary for the reduction of the affinity of polymerase for rRNA promoters.

Other types of mutant RNA polymerase should also exist. One would be insensitive to ppGpp and phenotypically relaxed. A second type would have an altered distribution of different forms of the enzyme with a consequent modified response to ppGpp *In vivo* such a mutant would have an unbalanced pattern of gene expression. No mutants of the first class have yet been isolated but one of the second class has been characterised.

This mutant, *ts*103[33], is unable to support the growth of lambdoid phages and has an altered pattern of protein synthesis. The isolated polymerase behaves as though the equilibrium between the high- and low-affinity rRNA forms is shifted in favour of the high-affinity isomer. The mutation maps at a locus distinct from those defining known subunits of RNA polymerase (Baralle *et al.*, manuscript in preparation).

An important consequence of promoter recognition by

Fig. 2 Multiple form model for RNA polymerase specificity. This is the simplest model assuming only two forms.
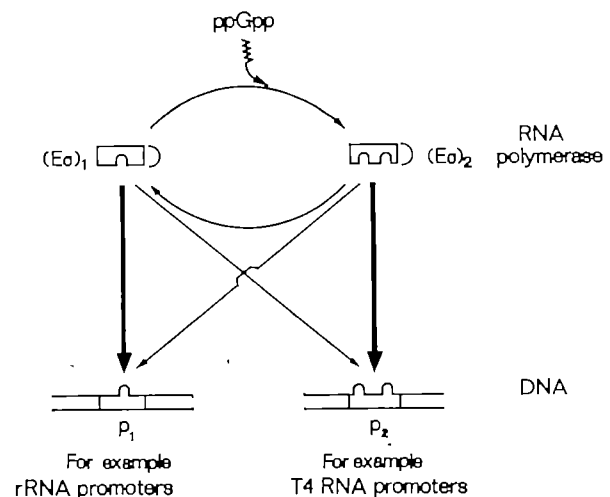
multiple forms of RNA polymerase is the possibility that the relative affinity of different forms of the enzyme for a promoter can vary. Thus at one extreme a promoter may bind tightly only one polymerase isomer, while in contrast another promoter may interact equally well with two or more forms of the enzyme. In this way a virtually continuous spectrum of recognition characteristics can be generated so that the promoter for any particular transcriptional unit is adapted to the control of polymerase specificity in a manner appropriate to the cell's requirements. Such a model predicts that the responses of the synthesis of particular RNA species to ppGpp will not fall into a small number of distinct classes but rather that the synthesis of each RNA species will have a characteristic individual response. The available evidence *in vitro* supports this view. Thus in a crude cell lysate the responses of the synthesis of four proteins, β-galactosidase, EF-Tu, β and β' subunits of RNA polymerase and ribosomal protein L7/L12 to ppGpp are quantitatively very different[34].

For simplicity I have so far considered RNA polymerase specificity formally in terms of two isomers. One indication that the regulation of promoter recognition by the enzyme may be more complex is the observation that in an *in vitro* system containing purified RNA polymerase the K_i for the preferential inhibition of su_3^+ tRNA synthesis from φ80 psu_3^+ DNA and of poly(G) synthesis from poly d(I-C) is 5 μM (refs 35 and 36). This contrasts with the value of 150 μM for rRNA synthesis and suggests that RNA polymerase may exist in at least three forms. Significantly, concentrations of ppGpp up to 10 μM which strongly inhibit su_3^+ tRNA synthesis substantially increase rRNA synthesis from φ80 d_3 DNA (Fig. 3).

Coupling of transcription and translation

I have argued that ppGpp alters polymerase specificity by converting one form of the enzyme to another. The rapid synthesis of this effector by an idling reaction on the ribosomes[12] is triggered by the codon-dependent binding of uncharged tRNA to the acceptor site[37,38]. The steady state concentration of the nucleotide is thus coupled to and requires a functioning ribosome. In addition, the rate of degradation of ppGpp is dependent on the energy input of the cell such that breakdown is facilitated when available energy is high[39,40]. In this way the level of ppGpp in the cell, and hence its effect on transcriptional specificity, is geared to both protein synthesis and energy metabolism.

Evidence *in vivo* shows that ppGpp, although a major

effector of transcriptional specificity, may not be the sole controlling element. The close involvement of transcription with translation is suggested by the suppression of ribosomal protein mutants by mutation of the β subunit of RNA polymerase[41]. Further, the antibiotics neomycin and spectinomycin, which inhibit polypeptide chain elongation, relieve the inhibitory effect of ppGpp on rRNA synthesis[42], an effect which is also observed in mutant bacteria with a thermolabile protein synthesis elongation factor G (ref. 43).

In vitro two components of the translation machinery have been shown to affect RNA polymerase specificity. One is fmet $tRNA_f^{met}$ whose specific binding to *E. coli* polymerase holoenzyme is dependent on both formylation and the initiator tRNA moiety[44]. This charged tRNA alters polymerase activity, in particular promoting transcription of $\lambda plac$ DNA and reducing rDNA transcription. A second effector is the elongation factor TuTs which decreases the transition temperature for rRNA synthesis on *E. coli* DNA while simultaneously increasing the sensitivity of transcription to ppGpp[15,30]. With φ80 d_3 DNA as template the factor sets rRNA synthesis at a level intermediate between the maximum level, observed at 5–10 μM ppGpp and the minimum level observed at 200 μM ppGpp (A.T., unpublished observations), and concomitantly increases su_3^+ tRNA synthesis[30]. Thus the pattern of transcription observed in the presence of EF-TuTs is that characteristic of the enzyme in a state which is maximally sensitive to ppGpp.

RNA polymerase and growth

The number of ribosomes a bacterium produces is in general approximately adjusted to the environment it grows in so that during exponential growth no more ribosomes are synthesised than can be engaged in protein synthesis. Thus the rate of protein production, and hence of growth, is balanced with the available resources. This correlation has led to the view that the level of ribosomes limits protein synthesis and thus control of ribosome production is a major facet of the control of growth[45]. At moderate to high growth rates the synthesis of the major transcriptional products required for ribosome biosynthesis, r-protein mRNA and rRNA, is coordinately regulated but at low growth rates this relation no longer holds, rRNA being relatively overproduced[46]. However, the synthesis of both products is under stringent control, both *in vivo*[47] and *in vitro*[34], suggesting that the mode of transcriptional regulation in the two cases, although not identical, is rather similar.

I have argued that an important aspect of the control of
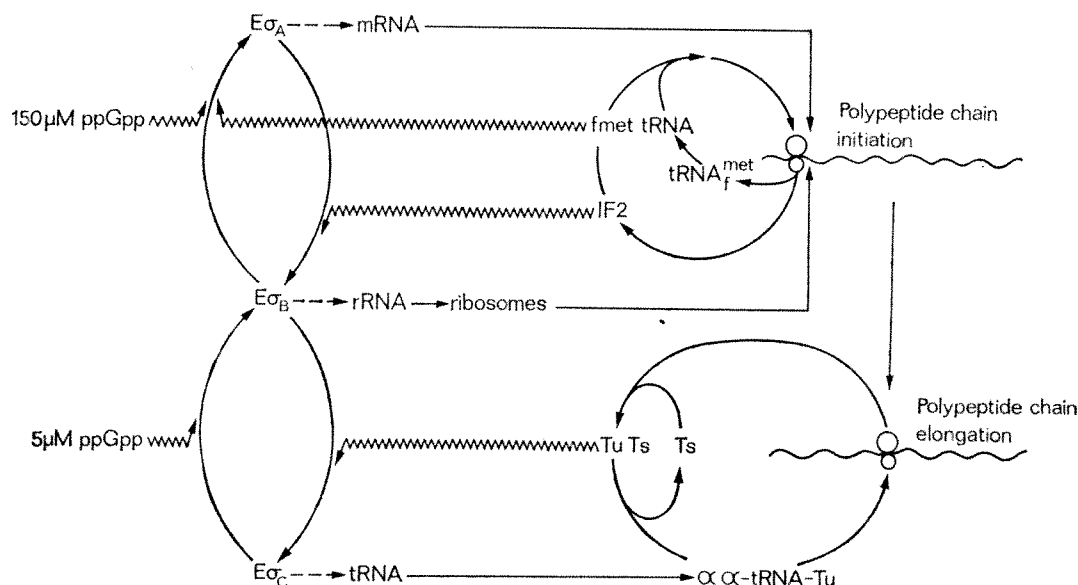
Fig. 3 A possible simplified scheme for the coupling of translation and transcriptional specificity. The types of RNA synthesised by the different forms of RNA polymerase represent the preferred specificity of the enzyme and do not represent absolute preferences. The term mRNA includes only certain mRNA species, for example, *lac* mRNA and T4 mRNA and excludes others, for example r-protein mRNA. The arrows indicating the direction of change in polymerase specificity do not imply any particular effecting mechanism.

rRNA synthesis is the response of RNA polymerase to intracellular signals. The question remains how such signals are integrated. Any molecular model for a control scheme is at present necessarily conjectural but must consider the logic of the control, the nature of the effectors and the dynamic balance between them.

One element in such a scheme is ppGpp acting as an effector of polymerase specificity in a manner similar to its *in vitro* function. Since the nucleotide probably alters the position of an equilibrium between specificity isomers the actual effect of ppGpp on transcription *in vivo* will depend on the relative free concentrations of ppGpp and of any effector which shifts the equilibrium in the reverse direction. Thus the mere accumulation of ppGpp need not invariably result in a shut off of stable RNA synthesis.

A second element is the alteration of the distribution between different specificity forms of RNA polymerase by the cycling factors of translation. Such effectors would signal the need for rRNA or for certain mRNA species or for even tRNA depending on which component of the translation machinery was momentarily limiting. For example free EF-TuTs might shift the polymerase towards tRNA production, assuming that the response of su_3^+ tRNA synthesis to effectors *in vitro* is representative of those tRNA species which are not cotranscribed with rRNA. Such a control constitutes a simple feedback loop since free EF-TuTs would be expected to accumulate when the supply of aminoacyl-tRNA was limited. Moreover the relatively high concentrations of EF-TuTs in the bacterial cell might be sufficient to override the low concentrations of ppGpp necessary to switch off su_3^+ tRNA synthesis by polymerase alone. Similarly fmet $tRNA_f$ could shift the pattern of transcription towards the synthesis of particular mRNA species. Again the logic is that of a simple feedback loop since fmet tRNA is consumed by the act of initiation of a polypeptide chain. The identity of the factor(s) maximising rRNA synthesis remains unknown[14]. It would seem logical, however, for rRNA production to be regulated by a factor involved in the initiation of protein synthesis. One candidate would be IF-2. The existence of several effectors regulating RNA synthesis implies that a mutation regulating in any one of the control elements would not eliminate entirely the synthesis of a particular RNA species but rather that such a change might only alter quantitatively the response to, for example, changes in growth medium. The number of bacterial proteins that interact with RNA polymerase is large[48] and thus the examples cited above possibly represent only a small fraction of the available control potential.

The data I have discussed relate to the *E. coli* holoenzyme in which the core polymerase is complexed with the σ subunit, a polypeptide of molecular weight ~90,000. In addition to σ, however, σ′, a polypeptide of molecular weight ~56,000 of apparently similar function, is sometimes found complexed with core enzyme in place of σ[**]. In *B. subtilis* RNA polymerase is isolated exclusively in this form[**] while in *E. coli* σ′ can be purified from ribosomes (R Buckland and A. Travers, unpublished). The relative roles of σ and σ′ *in vivo* are at present unknown and thus the possibility that the regulation of template selection by the σ′-containing enzyme differs from that of the σ-containing enzyme remains open.

The model outlined above envisages that the regulation of polymerase specificity is but one facet, albeit an important one, in the complex interplay of different modes of molecular control operating during normal bacterial growth. The transcribing enzyme complements the function of the highly specific DNA binding control elements for individual operons. In such a hierarchy, regulation of RNA polymerase specificity would be essentially a coarse control while fine control of particular transcripts would be provided by operon specific regulators.

Eukaryotic polymerases

To what extent are the principles of control of transcriptional specificity in prokaryotes applicable to eukaryotes? The transcriptional machinery of eukaryotes differs in at least one major respect from that of prokaryotes. In place of the single eukaryotic polymerase the nucleus of higher eukaryotes contains three structurally distinct enzyme types[**]. One, polymerase A, is found in the nucleolus and synthesises rRNA. A second, polymerase B, is located in the nucleoplasm where it synthesises hnRNA while the third, polymerase C, synthesises 5S and tRNA. These enzymes are all heteromultimers whose subunit structures, although more complex, bear strong resemblances to those of the bacterial polymerase. If the prokaryotic and eukaryotic enzymes are indeed homologues it would seem unlikely that the property of specificity modulation acquired by the prokaryotic polymerase would be entirely lost during the emergence of the eukaryotic line.

The separation of eukaryotic polymerases into three distinct structural types, each with apparently different functions, effectively releases the eukaryote from the prokaryotic restriction of balancing the use of a single enzyme. Such balancing is clearly necessary when the cell requires the simultaneous synthesis of two major classes of RNA, whose respective promoters are recognised by different conformations of the prokaryotic polymerase. The uncoupling in eukaryotes implies that each polymerase type would possess a relatively restricted initiation specificity, a view supported by the correlation of the template discrimination *in vitro* of the different enzymes with their presumed functions *in vivo*. The possibility thus remains that within the restricted range of promoter recognition of each enzyme, modulations of specificity could still result in the selective synthesis of different classes of RNA. In particular, for example, one polymerase might discriminate between promoters for the initiation of different mRNA species. A precedent for this exists in *E. coli* where the synthesis of certain mRNA species, for example those coding for ribosomal proteins and factors, is subject to stringent control[**] whereas the synthesis of other mRNA species, for example φ80 mRNA[**], is not. In eukaryotes such control might be particularly effective, for example, during the early stages of development.

If the eukaryotic polymerases are controlled what might their effectors be? Although ppGpp has been detected at low levels in cultures of *Dictyostelium*[**] this nucleotide has no apparent effect on the activity of fungal polymerases[**]. Other guanosine polyphosphates have, however, been isolated from organisms as diverse as the fungus, *Achyla*[**], and the crustacean *Artemia*[**,**]. The compounds are dinucleotides with guanosine residues connected by 5′–5′ polyphosphates. In addition, in some cases one of the 3′ positions is also phosphorylated with one or two phosphate residues. In *Achyla* the accumulation of these compounds is directly correlated with the rate of protein synthesis and thus they could arise by a capping of ppGpp, or more probably of its precursor pppGpp, in a manner analogous to the capping of mRNA[**]. Significantly these nucleotides strongly influence the activity of all three eukaryotic RNA polymerases. The response of the activity of these enzymes to increasing concentration of these nucleotides[**] bears strong similarities to the effect of ppGpp on the prokaryotic enzyme. The diguanosine nucleotides are thus candidates for one of the effectors of the RNA polymerases of lower eukaryotes Nevertheless, here, as also in higher eukaryotes, the biological importance of the control of RNA polymerase activity remains an open question.

continuing stimulating discussions, Drs B. Honda, G. Mitchison, O. Pongs and J. D. Smith for reading the manuscript and Drs M. Cashel and L. Soll for generously providing experimental materials.

1 Burgess, R. R., *J. biol. Chem.*, 244, 6168–6176 (1969).
2 Kuntzel, H., and Schafer, K. P., *Nature new Biol.*, 231, 265–269 (1975).
3 Gallant, J., and Lazzarini, R. A., in *Protein Synthesis. A Series of Advances* (edit. by McConkey, E. H.), (New York, Dekker, 1975).
4 Saucier, J. M., and Wang, J. C., *Nature new Biol.*, 239, 167–170 (1972).
5 Travers, A., Baillie, D. L., and Pedersen, S., *Nature new Biol.*, 243, 161–163 (1973).
6 Mangel, W. F., and Chamberlin, M., *J. biol. Chem.*, 249, 3002–3006 (1974).
7 Nakanishi, S., Adhya, S., Gottesman, M., and Pastan, I., *J. biol. Chem.*, 249, 4050–4056 (1974).
8 Richardson, J. P., *Proc. natn. Acad. Sci. U.S.A.*, 55, 1616–1623 (1966).
9 Bahr, W., Standen, W., Scheit, K. H., and Jovin, T. M., in *RNA polymerases* (edit. by Chamberlin, M. J., and Losick, R.), (Cold Spring Harbor Laboratory, New York, in the press).
10 Primakoff, P., and Berg, P., *Cold Spring Harb. Symp. quant. Biol.*, 35, 275–295 (1970).
11 Cashel, M., *J. biol. Chem.*, 244, 3133–3141 (1969).
12 Cashel, M., and Gallant, J., *Nature*, 221, 838–841 (1969).
13 Reiness, G., Yang, H. L., Zubay, G., and Cashel, M., *Proc. natn. Acad. Sci. U.S.A.*, 72, 2881–2885 (1975).
14 Block, R., in *Control of Ribosome Synthesis* (edit. by Maaloe, O., and Kjeldgaard, N. O.), 226 (Munksgaard, Copenhagen, 1976).
15 Travers, A., Kamen, R., and Cashel, M., *Cold Spring Harb. Symp. quant. Biol.*, 35, 415–418 (1970).
16 van Ooyen, A. J. J., de Boer, H. A., Ab, G., and Gruber, M., *Nature*, 254, 530–531 (1975).
17 Travers, A., *Molec. gen. Genet.*, 147, 225–232 (1976).
18 van Ooyen, A. J. J., Gruber, M., and Jorgensen, P., *Cell*, 8, 123–128 (1976).
19 Cashel, M., *Cold Spring Harb. Symp. quant. biol.*, 35, 407–413 (1970).
20 Nomura, M., Okomoto, K., and Asano, K., *J. molec. Biol.*, 4, 376–387 (1962).
21 Landy, A., and Spiegelman, S., *Biochemistry*, 7, 585–591 (1968).
22 Adesnik, M., and Levinthal, C., *J. molec. Biol.*, 48, 187–208 (1970).
23 Walter, G., Seifert, W., and Zillig, W., *Biochem. biophys. Res. Commun.*, 30, 240–247 (1968).
24 Goff, C. G., *J. biol. Chem.*, 249, 6181–6190 (1974).
25 Stevens, A., *Proc. natn. Acad. Sci. U.S.A.*, 69, 603–607 (1976).
26 Horvitz, H. R., *Nature new Biol.*, 244, 137–140 (1973).
27 Ratner, D., *J. molec. Biol.*, 89, 803–807 (1974).
28 Haseltine, W. A., *Nature*, 235, 329–333 (1972).
29 Baralle, F. E., and Travers, A., *Molec. gen. Genet.*, 147, 291–298 (1976).
30 Horvitz, H. R., *J. molec. Biol.*, 90, 739–750 (1974).
31 Travers, A., *Nature*, 244, 15–18 (1973).
32 Travers, A., *FEBS Lett.* (in the press).
33 Jacobson, A., and Gillespie, D., *Cold Spring Harb. Symp. quant. Biol.*, 35, 85–93 (1970).
34 Johnsen, M., and Fiil, N. P., in *Control of Ribosome Synthesis* (edit. by O. Maaloe and N. O. Kkeldgaard), 221 (Munksgaard, Copenhagen, 1976).
35 Debenham, P., and Travers, A., *Eur. J. Biochem.* (in the press).
36 Cashel, M., Hamel, E., Shapshak, P., and Boquet, M., in *Control of Ribosome Synthesis* (edit. by Maaloe, O., Kjeldgaard, N. O.), 279 (Munksgaard, Copenhagen, 1976).
37 Pedersen, F. S., Lund, E., and Kjeldgaard, N. O., *Nature new Biol.*, 243, 13–15 (1973).
38 Haseltine, W. A., and Block, R., *Proc. natn. Acad. U.S.A.*, 70, 1564–1568 (1973).
39 Friesen, J. D., Fiil, N. P., and von Meyenburg, K., *J. biol. Chem.*, 250, 304–309 (1975).
40 Hansen, M. T., Pato, M. L., Molin, S., Fiil, N. P., and von Meyenburg, K., *J. Bact.*, 122, 585–591 (1975).
41 Chakrabarti, S. L., and Gorini, L., *Proc. natn. Acad. Sci. U.S.A.*, 72, 2084–2087 (1975).
42 Muto, A., Kimura, A., and Osawa, S., *Molec. gen. Genet.*, 139, 321–327 (1975).
43 Kimura, A., Muto, A., and Osawa, S., *Molec. gen. Genet.*, 130, 203–214 (1974).
44 Pongs, O., and Ulbrich, N., *Proc. natn. Acad. Sci. U.S.A.*, 73, 3064–3067 (1976).
45 Maaloe, O., *Devl Biol. Supp.*, 3, 33–58 (1969).
46 Gausing, K., in *Control of Ribosome Synthesis* (edit. by Maaloe, O., and Kjeldgaard, N. O.), 292 (Munksgaard, Copenhagen, 1976).
47 Dennis, P. P., and Nomura, M., in *Control of Ribosome Synthesis* (edit. by Maaloe, O., and Kjeldgaard, N. O.), 304 (Munksgaard, Copenhagen, 1976).
48 Ratner, D., *J. molec. Biol.*, 88, 373–383 (1974).
49 Fukuda, R., Iwakura, Y., and Ishihama, A., *J. molec. Biol.*, 83, 353–367 (1974).
50 Losick, R., Sonenshein, A. L., Shorenstein, R. G., and Hussey, C., *Cold Spring Harb. Symp. quant. Biol.*, 35, 443–450 (1970).
51 Biswas, B. B., Ganguly, A., and Das, D., *Prog. Nucl. Acid Res. molec. Biol.*, 15, 145–184 (1975).
52 Klein, C., *FEBS Lett.*, 38, 149–152 (1974).
53 McNaughton, D. R., Klassen, G. R., and Lejohn, H. B., *Biochem. biophys. Res. Commun.*, 66, 468–474 (1975).
54 Lejohn, H. B., Cameron, L. E., McNaughton, D. R., and Klassen, G. R., *Biochem. biophys. Res. Commun.*, 66, 460–467 (1975).
55 Finamore, F. J., and Warner, A. H., *J. biol. Chem.*, 238, 344–348 (1963).
56 Warner, A. H., and Finamore, F. J., *Biochim. biophys. Acta*, 108, 525–530 (1965).
57 Ensinger, M. J., Martin, S. A., Paoletti, E., and Moss, B., *Proc. natn. Acad. Sci. U.S.A.*, 71, 3465–3469 (1975).

articles

Perception of melodies

H. C. Longuet-Higgins

Centre for Research on Perception and Cognition, Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, UK

A computer program has been written which will transcribe a live performance of a classical melody into the equivalent of standard musical notation. It is intended to embody, in computational form, a psychological theory of how Western musicians perceive the rhythmic and tonal relationships between the notes of such melodies.

A SEARCHING test of practical musicianship is the 'aural test' in which the subject is required to write down, in standard, musical notation, a melody which he has never heard before. His transcription is not to be construed as a detailed record of the actual performance, which will inevitably be more or less out of time and out of tune, but as an indication of the rhythmic and tonal relations between the individual notes. How the musical listener perceives these relationships is a matter of some interest to the cognitive psychologist. In this paper I outline a theory of the perception of classical Western melodies, and describe a computer program, based on the theory, which displays, as best it can, the rhythmic and tonal relationships between the notes of a melody as played by a human performer on an organ console.

The basic premise of the theory is that in perceiving a melody the listener builds a conceptual structure representing the rhythmic groupings of the notes and the musical intervals between them. It is this structure which he commits to memory, and which subsequently enables him to recognise the tune, and

to reproduce it in sound or in writing if he happens to be a skilled musician. A second premise is that much can be learned about the structural relationships in any ordinary piece of music from a study of its orthographic representation. Take, for example, the musical cliché notated in Fig. 1.

Fig. 1
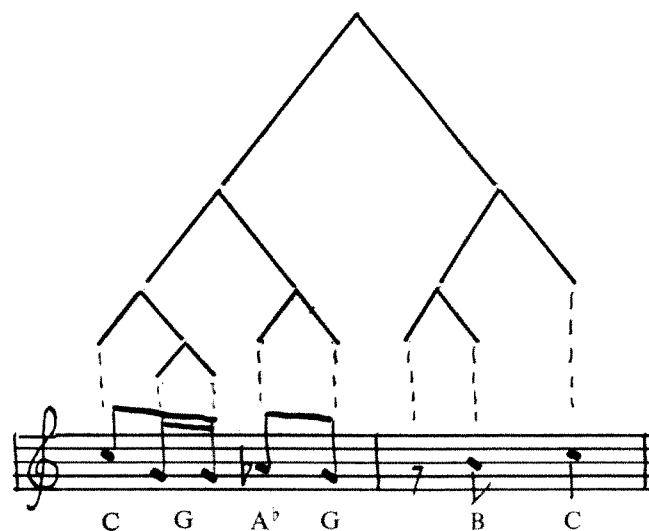
The way in which the notes are rhythmically grouped is evident from the disposition of the bar lines and the "beams" linking the notes of the first bar. The rhythm is, in this case, a binary tree each terminal of which is a note or a rest, but more generally, such a tree may have ternary as well as binary nodes.

The tonal relations between the notes in Fig. 1 are also indicated by the symbolism, but more subtly. It is a common mistake to suppose that the position of a note on the five-line stave (and its prefix, if any) indicates merely the approximate pitch of the note—where it would be located on the keyboard. If that were true, an equally acceptable alternative to Figure 1 would be Fig. 2, in which the A♭ has been written as a G♯, with the same location on the keyboard. But a music student

Fig. 2

who offered Fig. 2 as his transcription would lose marks for having misrepresented the tonal relation of the fourth note to its neighbours (though he could hardly be imagined not to have perceived it properly!).

The problems posed by melodic perception are not dissimilar to those which arise in the perception of speech. The distinction between the A♭ in Fig. 1 and the G♯ in Fig. 2 is analogous to the difference between the homophones "here" and "hear" in English; though these words sound exactly alike they are interpreted and spelt quite differently according to the context in which they are heard. Another problem in speech perception, which has its counterpart in the perception of melody, relates to the timing of successive acoustic events. The way in which the syllables of a poem are perceptually grouped into "feet" is largely unaffected by variations in rate of delivery, and the same applies to the rhythmic grouping of the notes of a melody. Notes which, on paper, are of equal length, will in a live performance be sounded at quite unequal intervals of time, particularly in an "expressive" performance. A change of metre from duplets to triplets can, nevertheless, usually be distinguished quite clearly from a mere quickening of tempo, in a reasonably competent performance. Previous programs for the automatic transcription of music have required the performer to maintain a fairly constant tempo[1,2]; but human listeners have no difficulty in discerning the rhythms of melodies played by performers who are free from this constraint.

The third premise of the theory is that the perception of rhythm and the perception of tonal relationships can be viewed as independent processes. This strong claim (which is not to be misunderstood as referring to the process of musical composition) may be weakly supported by two observations. First, that a given melodic sequence such as the ascending major scale will be heard as such by a Western musician regardless of the rhythm in which it is played. And conversely, that a "dotted" rhythm, for example, will be clearly recognisable for what it is, regardless of the musical intervals between the successive notes. To say this is not, of course, to deny that higher cognitive processes can and will operate on the "surface structure" generated by rhythmic and tonal perception, to reveal musically significant relations between the rhythm and the tonality. But one may reasonably suppose that such processes of musical appreciation can only begin when some structure has been created on which they can get to work.

Rhythm

One might imagine that to discern the rhythm of a melody the listener must be able to perceive differences in loudness between successive notes. This may be true on occasion but fails as a generalisation for two reasons. First of all, performers do not as a rule thump out every note which occurs on a beat or at the beginning of a bar; to do so would be as tiresome as to accent, in reading a poem, every syllable that occurred at the beginning of a foot. But more decisively, there are instruments such as the organ and harpsichord on which it is physically impossible for the performer to vary the acoustic intensity of each individual note; all he can control is the time of onset of the note and its temporal duration. It is nevertheless quite possible for a listener to perceive correctly the rhythm of a melody played on such an instrument; we conclude that temporal information alone is enough for the purpose, except in special circumstances.

The basic assumption underlying the rhythmic component of the program is that the first necessity in perceiving the rhythmic structure of a melody is to identify the time of occurrence of each "beat". Music in which the beat is irregular falls outside the scope of the theory, which therefore has nothing to say about the rhythmic perception of recitative or of music in which, for example, the beats alternate in length. The grouping of the beats into higher metrical units such as "bars" raises issues which have been discussed elsewhere[3,4]; the principal concern of the present study is with the manner in which each beat should be subdivided, and with the problem of keeping track of the beat through unforeseen changes in tempo.

In Western music by far the commonest subdivisions of the beat are into 2 and into 3 shorter metrical units; these in turn can be further subdivided into 2 or 3. Whether a beat, or a fraction of a beat, is perceived to be divided, depends, according to the theory, on whether or not it is interrupted by the onset of a note. What counts as an "interruption" is a matter of some delicacy, to which I shall return in a moment.

After such a process of division, and subdivision, every note will find itself at the beginning of an uninterrupted metrical unit. It is the relations between these metrical units which constitute the rhythm of the tune; the metrical units can be thought of as the nodes of a "tree" in which each non-terminal node has either two or three descendants. Every terminal node in the tree will eventually be attached either to a rest or to a note (which may be sounded or tied) in the manner of Fig. 1. The program does not actually draw such trees, nor print out a musical stave; it represents the rhythm in a nested bracketed notation. It also indicates the phrasing; if the offset of a note occurs earlier than half-way through its allotted time, or else appreciably before the end of that time, the note is marked "*stc*", standing for "*staccato*", or "*ten*" for "*tenuto*".

We now return to the question: what counts as an interruption? By what criterion could a listener judge whether the onset of a note occurs "during" the current metrical unit rather than "at" its beginning or its end? Plainly there must be some upper limit to the temporal discrepancies he can disregard, just as there is a lower limit to those he can detect. The former limit— the listener's "tolerance"—must obviously exceed the latter. It must be small enough to permit the structuring of rapid rhythmic figures, but large enough to allow a reasonable degree of flexibility to the tempo. In the program the tolerance can be preset to any desired value. Experiments with the program indicate that for reasonably careful performances a tolerance of about 10 cs meets both criteria, but that for more "expressive" performances, of relatively sedate melodies, a greater tolerance is needed if an obvious *rubato* is not to be misconstrued as a variation in rhythm.

In order to perceive the rhythm clearly a listener must, it is assumed, take account of the precise onset time of every note within any metrical unit in predicting when the unit could end. The rule eventually adopted for making such predictions was as follows: if, in the course of a binary metrical unit, a note which terminates the first subunit begins a little less than half way through, then the expected further duration of the unit is reduced in magnitude, to the mean of its original value and the value implied by the time of onset of the note in question.

$p =$ -7   -6   -4   -2   1   3   6   8   10   13

$p =$ -8   -7   -5   -3   -1   0   2   4   5   7   9   11   12
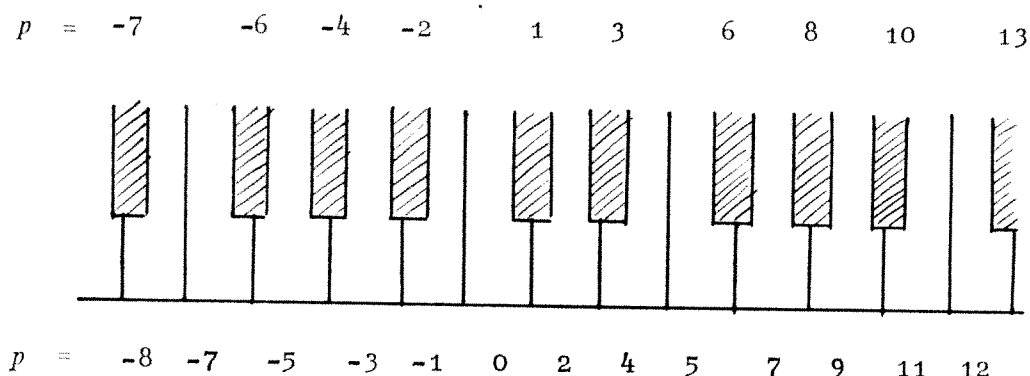
Fig. 3

Corresponding remarks apply, of course, in cases where the note is slightly late or the current metrical hypothesis assigns a ternary rather than a binary structure to the metrical unit in question. In fact the program also allows for the termination of a metrical unit, not by the actual onset of a note, but by the anticipated end of a lower metrical unit which is itself interrupted by the onset of one or more notes. Such procedures are, unfortunately, much more difficult to specify precisely in English than in a suitably designed programming language; but this fact only underlines the value of casting perceptual theories in computational form.

Finally, it is necessary to commit oneself, in writing such a program, to a view as to what counts as good perceptual evidence for a change in metre. It is here assumed that the listener initially expects a pure binary metre, but is prepared to change his mind at any level in the metrical hierarchy. The evidence for a change in metre may be of two kinds: that the current metre implies a "syncopation", in which the beginning of the next beat, or higher metrical unit, is not accompanied by the onset of a note; or that it implies a "distortion" in which an excessively large change of tempo is required to accommodate the current metrical hypothesis. Each of these outcomes represents a flouting of the listener's expectations and either may, according to the theory, lead him to change his opinion about the metre if the other possible division of the current metrical unit (ternary instead of binary, or vice versa) does not imply a distortion or a syncopation. Lastly, it is assumed that once having changed his mind the listener does not change it back again until he encounters positive evidence for doing so.

Tonality

In committing a melody to memory the listener must not only create a rhythmic structure of the kind depicted in Fig. 1; he must also identify the tonality of each note which is to be attached to it. This tonal information should ideally suffice, not only for the transcription of the melody into standard notation,

but also for the purpose of evaluating the intonation of a performance on, say, the violin, which permits fine distinctions of pitch which cannot be made explicit on a keyboard instrument (see Fig. 3.)

To appreciate what is involved in this task it is necessary to formalise the classical theory of tonality, as developed by Rameau[5], Bosanquet[6], Helmholtz[7] and other writers. In the formal theory[8] every musical note is assigned coordinates (x, y, z) in a "tonal space" of three dimensions, corresponding to the perfect fifth, the major third and the octave, respectively. (The ideal frequency ratios of these intervals are 3/2, 5/4 and 2/1 respectively—involving the first three prime numbers—so that they are strictly incommensurable when not distorted by equal temperament.) Thus in Fig. 4, if the origin is taken as "middle C", the tonal coordinates of the various notes are as shown in the first three rows of the table. The following points may be noted: (1) It is the relative values of the coordinates (x, y, z), not their absolute values, which characterise the melodic sequence. Thus, increasing all the z values by 1 would merely put the melody up one octave. (2) The numerical values of the coordinates (x, y, z) are all small; this is evidently the result of having chosen middle C as origin. (3) The "position" p of each note on the keyboard, that is, its distance above middle C in keyboard semitones, is given by

$$p = 7x + 4y + 12z$$

so that there are arbitrarily many different notes (x, y, z) with the same value of p. (4) The conventional name of any note is determined not by its keyboard position p but by its "sharpness" q, defined as

$$q = x + 4y$$

The name N of the note is such that there are q sharps, or $-q$ flats, in the key signature of N major. Thus an "A" is a note with $q = 3$, and an "A♭" is one with $q = -4$.

x	0	1	0	1	.	1	0	(0)
y	0	0	-1	0	.	1	0	(2)
x	1	0	1	0	.	0	1	(0)
p	12	7	8	7	.	11	12	(8)
q	0	1	-4	1	.	5	0	(8)
N	C	G	A	G	.	B	C	(G)

Fig. 4
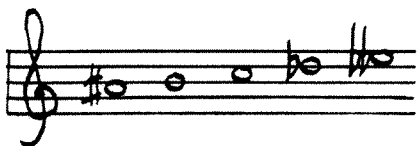
Fig. 5

The task of naming the notes of a melody played on the keyboard, or of notating them correctly on the five-line stave, therefore involves the apparently insoluble problem of determining the three coordinates (x, y, z) of each note from its

taining chromatic scales, because each keyboard semitone in such a scale would be assigned the same value of δq, and this would lead to absurd tonal interpretations such as that shown in Fig. 5.

An alternative, and musically much more plausible, hypothesis is that the listener identifies the sharpness of each note by placing it within a diatonic interval of the very first note (or perhaps an interval of degree 6 if the span indicates that the interval is diabolic). Such a rule would account equally well for the q values of the notes in Fig. 4. It would, furthermore, account very nicely for the sharpness of the notes in the theme of Bach's Musical Offering (Fig. 6)

| p | 0 | 3 | 7 | 8 | -1 | 7 | 6 | 5 | 4 | 3 | 2 | 1 | 0 | -1 | -3 | -5 | 0 | 5 | 3 | 2 | 0 |
| q | 0 | -3 | 1 | -4 | 5 | 1 | 6 | -1 | 4 | -3 | 2 | -5 | 0 | 5 | 3 | 1 | 0 | -1 | -3 | 2 | 0 |
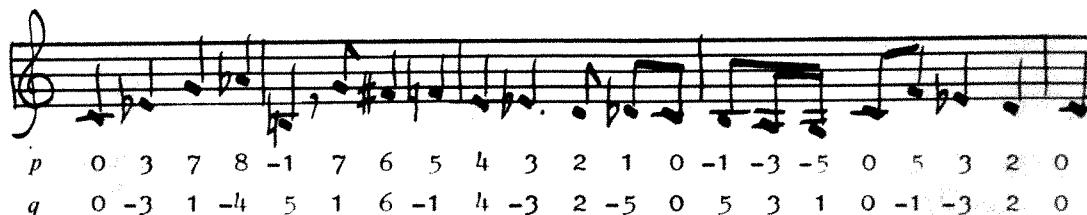
Fig. 6

keyboard position p, so as to be able to determine its sharpness q. The problem is analogous to the visual problem of adding an extra dimension to the 2-dimensional image of a three-dimensional scene, except that a keyboard performance of a melody supplies the listener with only one directly audible dimension, to which he must add two more to identify each note uniquely.

There is in fact a short cut to the solution of this problem, arising from the mathematical fact that a given choice of p severely restricts the range of possible values of q. A little simple arithmetic shows that q must differ from $7p$ by a multiple of 12. If, therefore, we can find independent grounds for limiting q to one of a fairly small set of values, we can determine it uniquely from the remainder upon division of p by 12. A survey of the published scores of classical melodies reveals[6] that the value of q (which can be directly determined from the notation) never changes from one note to the next by more than 11 units of sharpness. As a consequence, if δp is the "span" of any interval between two successive notes, then the "degree" δq of that interval is restricted to the following alternative values when δp lies in the range -11 to $+12$

$$\delta p = \begin{array}{cccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \text{or} & -11 & -10 & -9 & -8 & -7 & -6 & -5 & -4 & -3 & -2 & -1 & 0 \end{array}$$

implies that

$$\delta q = \begin{array}{cccccccccccc} -5 & 2 & -3 & 4 & -1 & 6 & 1 & -4 & 3 & -2 & 5 & 0 \\ \text{or} & 7 & -10 & 9 & -8 & 11 & -6 & -11 & 8 & -9 & 10 & -7 \end{array}$$

By far the commonest intervals occurring in classical and traditional melodies are "diatonic" intervals, with $|\delta q| < 6$. "Chromatic" intervals, with $|\delta q| > 6$, and "diabolic" intervals, with $|\delta q| = 6$, are relatively rare. If they were nonexistent, the degree δq of each interval would be uniquely determined by its span δp, and one could infer the sharpness of each note in any melody from its position relative to its immediate predecessor. Such a "Markovian" theory of tonal perception would, as it happens, correctly predict the q values in Fig. 4. It would, however, fail dismally for melodies con-

In this melody there are in fact four chromatic intervals—a diminished seventh between the 4th and 5th notes, of degree $\delta q = -9$, and three chromatic semitones, of degree -7. Such intervals could not, for obvious reasons, be correctly identified by the "Markovian" procedure.

It will not do, however, to assume that the first note is invariably the "keynote" to which every other note should be referred in order to determine its sharpness. First because melodies very often begin on notes other than the keynote, or "tonic", and second because the tonic may very well seem to change in the course of a melody, when we speak of a "modulation" having occurred.

A good example of an indisputable modulation is to be found in the subject of the B minor fugue in the first book of Bach's *Wohltemperierte Klavier* (see Fig. 7). The first three notes clearly establish the tonic as B ($q = 5$), and all the other notes in bars 1 and 2 are related to it by non-chromatic intervals ($|\delta q| < 7$). But the first note of bar 3, though in the same keyboard position as the C in bar 2, is notated differently. To have written it as a C would have produced the sequence C♯ C C♯, calling for two chromatic semitones in succession. But in unaccompanied classical melodies such an event never seems to occur, for the very good reason that if $X Y Z$ are three successive notes of a melody which, on paper, are separated by chromatic intervals $X Y$ and $Y Z$, then there is always an alternative, simpler, interpretation of the middle note Y which transforms both intervals into diatonic ones. Generally speaking, then, the tonal identity of a note cannot be finally established until the following note is heard. In Fig. 7 the offending note has become transformed into a B♯, making both the neighbouring intervals into diatonic semitones, of degree 5 and -5 rather than -7 and 7 respectively. But a B♯ is too far from the old tonic B to belong to its key, so that a modulation is perceived to occur to a new key, that of F♯, such that the value of δq for the new note B♯ is only 6, which is just close enough for comfort.

There seem to be other general restrictions upon the contexts in which chromatic intervals occur in classical melodies. The

Fig. 7

```
: printlist(cliche);
```

```
[ 12 154 227]
[ 36 285 294]
[ 31 322 327]
[ 31 336 341]     a
[ 32 349 383]
[ 31 384 407]
[ 35 445 453]
[ 36 484 527]
```

Fig. 8

```
: 10->tolerance; notate(cliche);                                    b
```

```
[[[ 24 C STC] [[-5 G STC] [ 0 G STC]]] [[ 1 AB] [-1 G TEN]]]

[[[REST] [ 4 B STC]] [ 1 C TEN]]
```

```
:
```


c
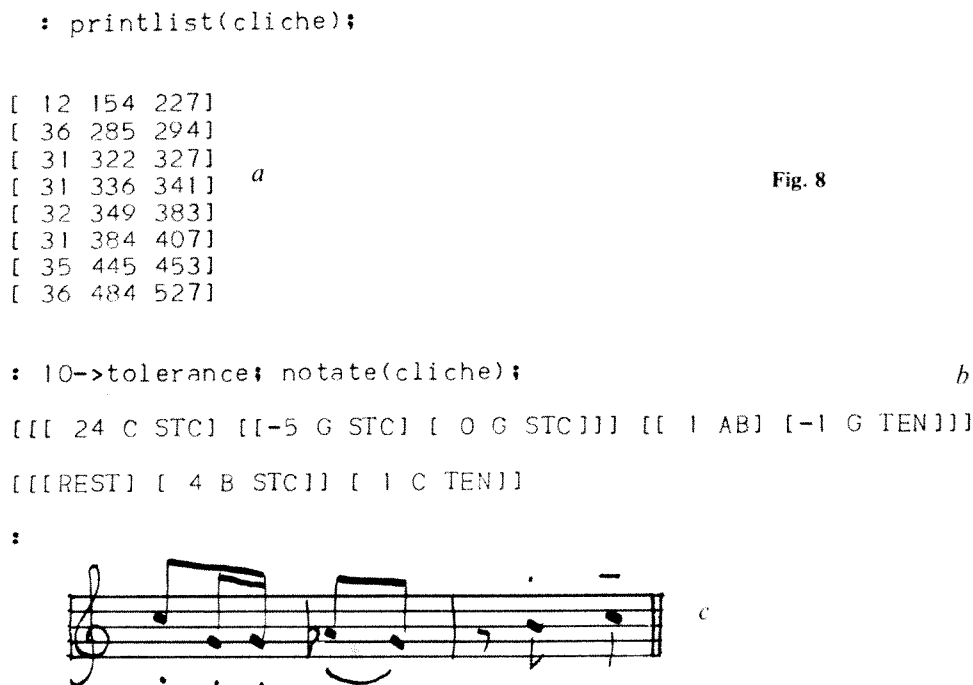
most important of these relates to four-note sequences *W X Y Z* in which the middle interval is chromatic. In such a sequence not only must both *W X* and *Y Z* be non-chromatic, but at least one of the intervals *W Y* and *X Z* must be diatonic. If the interpretations of *W*, *X*, *Y* and *Z* based on the current key violate this rule, then the tonality of the note *Y* is reinterpreted in such a way as to make *X Y* a diatonic interval, and to force a modulation into a key to which *Y* belongs. As implied by what has been said, a note is regarded as belonging to a given key if its sharpness relative to the tonic lies in the range −5 to +6 inclusive.

Another rule which seems to be necessary in order to account for the notation of chromatic scales, particularly in music of the period following Bach, concerns the tonal interpretation of ascending semitones. If such a semitone ends on the note of a key whose sharpness relative to the tonic is 2, 3, 4 or 5, then its first note is to be assigned a relative sharpness of 7, 8, 9 or 10. Though this reassignment places the note outside the key, it does not by itself precipitate a modulation; if it did, then an ascending chromatic scale of any length would trigger a whole sequence of modulations into progressively sharper and sharper keys.

Two further rules are necessary, and sufficient, for determining the relative sharpnesses of the notes in most classical melodies. The first is a rule to the effect that for the purposes of establishing tonality one may conflate repeated notes, or notes separated by an octave (the second and third notes in Fig. 4 provide an example). The other rule, which is theoretically less satisfactory, is that the tonic may be determined from the first two notes, and that it will be either the first note itself or the note a fifth below it. This rule, and the absence of any more delicate tests of modulation than those already described, are undoubtedly the weakest links in the tonal section of the program.

Before I describe the program in detail, a few words of caution may be in order. First, the tonal rules outlined above must not be expected to apply to accompanied melodies, where the accompaniment supplies tonal information which may not be implicit in the melody itself. Nor must it be supposed that the rules necessarily hold for covertly polyphonic melodies in which, for example, alternate notes really belong to two different melodies. Further, the contextual constraints on chromatic intervals will often be violated at phrase boundaries, marked by rhythmically prominent rests, though this is not

always the case. And finally, one must allow for the possibility that in a musical score a radical change of notation (such as occurs between the first and second sections of Chopin's *Raindrop Prelude*) does not signify a real change in tonality, but merely an "enharmonic change" designed to simplify the reader's task. Only if such qualifications are borne in mind can the program safely be used to indicate how a melody performed on the keyboard should be transcribed into conventional notation.

The program

The program accepts as input a list of sublists, each of which comprises three numbers. The first number is the keyboard position of the corresponding note and lies in the range 0 to 48, there being four octaves on the organ console. The second number is the time in centiseconds at which the note was depressed, and the third number indicates the time at which the note was released. The order of the notes on the list is the order of their times of onset. The list itself is generated from a live performance of a melody on an electronic organ connected, through an analogue-to-digital converter, to a high speed paper tape punch. The information on the paper tape is equivalent to the information which would be recorded on a player piano roll, and no more. The preprocessing of the paper tape is an entirely automatic matter, which simply involves constructing the above-mentioned list from the paper tape record and transferring it to disk storage.

The performer is required, by the present version of the program, to establish the initial tempo and the number of beats in a bar by prefacing his performance of the melody by a bar's worth of beats on some low note, which may conveniently be positioned an octave below the first note of the melody, so as not to prejudice the tonality.

The program itself is written in POP2, the high level programming language designed and developed in Edinburgh by Burstall and Popplestone[9]. It is relatively short, and is structured as follows. First, the list of sublists is converted into a list of records, each of which has a "slot" indicating the pitch, onset time and offset time of a particular note, and further slots which are to hold the span δp and the degree δq of the interval between the note and its predecessor. The keynote is then fixed by the positions of the first two notes, and the relative sharpnesses of all the notes are determined from their keyboard positions by an algorithm based on the theory of tonality

```
: printlist(tris);

[ 12  24  .114]
[ 12 148  238]
[ 24 274  399]
[ 31 400  554]
[ 34 551  587]
[ 32 586  671]
[ 27 669  7.11]
[ 32 707  794]
[ 26 795  831]
[ 31 829  860]
[ 24 863  895]
[ 29 895  989]
[ 31 987  1021]
[ 29 1020 1145]
[ 27 1140 1242]
[ 26 1268 1282]
[ 24 1289 1298]
[ 22 1308 1320]
[ 29 1332 1452]
[ 26 1450 1495]
[ 22 1508 1517]
[ 21 1528 1536]
[ 20 1546 1556]
[ 27 1570 1696]
[ 24 1692 1734]
[ 20 1752 1762]
[ 19 1774 1782]
[ 18 1792 1808]
[ 26 1815 1930]
[ 29 1928 1934]
[ 27 1932 2062]
[ 26 2059 2188]
[ 25 2183 2446]
[ 24 2491 2628]
```

Fig. 9

a

outlined in the previous section. The next stage is a rhythmic analysis (which could have been carried out first, as it is indifferent to the results of the tonal routines). Each beat is examined in turn, by a combination of "top down" and "bottom up" analysis in which the time of onset of each note is used both for establishing the structure of the rhythmic hierarchy and for correcting the estimated tempo. In the course of this analysis the time of offset of each note is used for determining how the note was phrased.

The final stage in the operation of the program corresponds to the exercise of musical literacy; it consists of displaying, on paper, the essential features of the structure created by the rhythmic and tonal analyses as a sequence of nested lists of symbols. The innermost symbols name the individual notes as, for example, 'D' (D natural), 'DS' (D sharp), or 'DB' (D flat); the word REST is self-explanatory. Each name is preceded by either the word TIED if the note is tied to its predecessor or a number indicating the span (not the degree, which is implicit in the name of the note) of the interval from the preceding note; this is needed for identifying the octave in which the note occurs. Finally, a note which is not tied may be followed by the abbreviation STC or TEN indicating that the note was played *staccato* or *tenuto*; the absence of either abbreviation implies that the note was played *legato*.

Figures 8, 9 and 10 provide examples of the program's performance. Each figure indicates (*a*) the "raw" input, in which each set of three numbers gives the keyboard position of a note and its times of onset and offset in centiseconds, (*b*) the output generated by the program from the input (*a*), and (*c*) the result of transcribing the output (*b*) by hand into ordinary stave notation.

The performance of the tune shown in Fig. 8 was prefaced by a single low C, and the time between the onset of this note and the next was arbitrarily taken as a "minim" in adopting the note values indicated in (*c*); it will be noted that in (*b*) the outermost brackets enclose a minim's worth of notes. The interpretation (*b*) was obtained from the input (*a*) with a tolerance of 10 cs; with a tolerance of 15 cs the program would assign the two semiquavers to the same node. It is worth noting that the actual times of onset of the first four quaver

```
: 13->tolerance; notate(tris);

[ 12 C][ 7 G]

[[[TIED G] [ 3 BB]] [-2 AB]][[[TIED AB] [-5 EB]] [ 5 AB]]

[[[TIED AB] [-6 D]] [[ 5 G] [-7 C]]][[ 5 F] [[TIED F] [ 2 G]]]

[-2 F][-2 EB TEN]

[[[-1 D] [-2 C STC] [-2 BB]] [ 7 F]][[TIED F] [-3 D TEN]]

[[[-4 BB STC] [-1 A STC] [-1 AB]] [ 7 EB]][[TIED EB] [-3 C TEN]]

[[[-4 AB STC] [-1 G STC] [-1 FS]] [ 8 D]][[TIED D] [ 3 F -2 EB]]

[[TIED EB] [-1 D]][[TIED D] [-1 DB]]

[TIED DB][TIED DB]

[-1 C]
```

b

c

```
: printlist(stan);
```

```
[ 19  148  190]
[ 31  280  287]
[ 29  302  309]
[ 27  322  329]
[ 34  347  466]
[ 31  474  518]
[ 27  538  548]
[ 26  559  566]
[ 25  578  586]
[ 33  605  648]
[ 30  646  657]
[ 26  669  678]
[ 25  687  696]
[ 24  707  714]
[ 32  729  760]
[ 29  769  777]
[ 25  791  801]
[ 24  811  820]
[ 23  830  839]
[ 31  856  987]
[ 27  986  1027]
[ 24  1049  1054]
[ 23  1068  1075]
[ 22  1087  1096]
[ 29  1111  1153]
[ 26  1152  1157]
[ 22  1174  1183]
[ 21  1194  1202]
[ 20  1211  1220]
[ 27  1232  1270]
[ 24  1272  1279]
[ 20  1295  1304]
[ 19  1316  1325]
[ 18  1336  1348]
[ 26  1360  1619]
```

Fig. 10

a

units differed in the performance by 37, 27 and 35 cs respectively, the separation between the last two notes being 39 cs. The considerable discrepancy between these numbers clearly illustrates the acute difficulty which would confront any attempt to determine the rhythm without taking account of its hierarchical structure.

Figure 9 shows how the program handled a performance of part of the long cor anglais solo from the prelude to Act III of Wagner's *Tristan und Isolde*. This example is interesting in two particular respects. First, it involves the perception of a change from a binary to a ternary metre in the fifth bar; and secondly, the published score indicates a grace note at the end of the seventh bar, to which it would be inappropriate to assign a separate place in the rhythmic structure. The program's output agrees fully with the score in its rhythmic and tonal indications; there are slight discrepancies in the marks of phrasing—Wagner marked all the triplet quavers as staccato—but for this the performer is clearly to blame, not the program.

Finally, Fig. 10 illustrates the program's handling of a later section of the same melody (prefaced, this time, by only one "cue" note, which is why each line of the output contains only one, not two, minim units). Again the rhythm is correctly represented, though there are minor discrepancies in phrasing. As for the tonality, the main point to note is the perceptual problem presented by the rapid succession of modulations beginning in the second bar. There is, nevertheless, only one note to which the program assigns a tonality at variance with that indicated by Wagner: he wrote the second C♯ in the second bar as a D♭.

Conclusions

The domain of competence of the program is, of course, very restricted: it cannot be expected to reveal significant tonal or rhythmic relations between the notes of "atonal" or "arhythmic" melodies, for example. But the perceptual theory on which it is based does seem worthy of serious consideration, in that up to the present time no detailed suggestions seem to have

```
: 13->tolerance; notate(stan);

[[[ 12 G STC] [-2 F STC] [-2 EB STC]] [ 7 BB]]

[[TIED BB] [-3 G TEN]]

[[[-4 EB STC] [-1 D STC] [-1 CS STC]] [[ 8 A] [TIED A] [-3 FS]]]

[[[-4 D STC] [-1 CS STC] [-1 C STC]] [[ 8 AB] [REST] [-3 F STC]]]

[[[-4 DB STC] [-1 C STC] [-1 B STC]] [ 8 G]]

[[TIED G] [-4 EB TEN]]

[[[[-3 C STC] [-1 B STC]] [-1 BB STC]] [[ 7 F] [TIED F] [-3 D STC]]]

[[[-4 BB STC] [-1 A STC] [-1 AB STC]] [[ 7 EB] [TIED EB] [-3 C STC]]]

[[[-4 AB STC] [-1 G STC] [-1 FS]] [ 8 D]]
```

b

c

been offered as to how a listener builds an internal representation of a melody from a live performance. The most significant rhythmic hypothesis in the theory is that the rhythm of a melody is conceptualised as a structural hierarchy, and that the onset of each note provides important predictive information about the time of onset of the following note at every level in the hierarchy. The hypotheses underlying the tonal section of the program are presumably limited in application to the kind of music that has been developed in the West; but for such music one conclusion at least seems secure, namely that the tonality of any note cannot in general be established unambiguously until the following note has been heard. It is perhaps surprising that such a limited amount of context should usually suffice for the purpose, but it should be remembered that it is really the key of the melody which creates the tonal context in the first place.

It seems altogether possible that the principles of operation of the program's rhythmic component will apply to other temporal processes such as the perception of speech.

A full account of the program and of the underlying theory is to be published elsewhere.

I thank D. C. Jeffrey, M. J. Steedman, B. C. Styles, O. P. Buneman and G. E. Hinton for practical assistance and helpful discussions, and to the Royal Society and the SRC for research support.

1 Styles, B C., thesis, Cambridge Univ. (1973)
2 Askenfelt, A , *Quarterly Progress and Status Report, Speech Transmission Laboratory, Royal Institute of Technology, Stockholm,* 1, 1 (1976)
3 Longuet-Higgins, H. C , and Steedman, M. J., *Machine Intelligence,* 6, 221 (1971).
4 Steedman, M. J , thesis, Univ Edinburgh (1973).
5 Rameau, M , *Traité de l'harmonie reduite à des principes naturels* (1721)
6 Bosanquet, P. H M., *Elementary Treatise on Musical Intervals and Temperament* (1876)
7 Helmholtz, H L. F (Ellis's translation), *On the Sensations of Tone,* 2nd ed (London, 1885)
8 Longuet-Higgins, H C., *Music Review,* 224 and 271 (1962), *Proc R. Inst ,* 45, 87 (1972).
9 Burstall, R. M , Collins, J. S., and Popplestone, R J , *Programming in POP-2* (Edinburgh University Press, Edinburgh, 1971).

New evidence on the origin of rotation measures in extragalactic radio sources

P. P. Kronberg*

Max-Planck-Institut für Radioastronomie, Bonn

Martine Simard-Normandin

David Dunlap Observatory, University of Toronto

The causes of the Faraday rotation in extragalactic radio sources have been reinvestigated in the light of new polarisation data which extend to large redshifts. Our results show convincingly that at nearly all galactic latitudes, the large rotation measures are generated outside of the Milky Way. Furthermore, any contribution from the intergalactic medium must be small when compared with the Faraday rotation generated at or in the extragalactic sources themselves. Having rotation measures for 108 QSOs of which 48 have redshift (z) greater than unity, we have been able to place the lowest and most reliable upper limits on any contribution from an intergalactic magneto-ionic medium at early cosmological epochs. The data show promise that it should soon be possible to ascertain whether or not the density of intergalactic ionised gas follows the cosmological scale factor.

THE Faraday rotation of radio waves from extragalactic sources has three possible origins: (1) a 'local' contribution in the Milky Way; (2) a possible contribution from intervening intergalactic material and (3) a contribution from the sources themselves. (We shall ignore ionospheric Faraday rotation, which is very small and can easily be removed.) The fact that there is a galactic contribution is well established, since the rotation measures (RM) show systematic variations when plotted on galactic co-ordinates[1-4]. Less clear, however, is whether the effects (2) and (3) are discernable, and there are conflicting claims in the literature[3-8] as to whether any inter-

galactic component can be seen. The source component (3) has hitherto been thought to be small[1-3].

In an effort to improve the situation we have measured linear polarisations at one or more wavelengths for a large number of QSOs and radio galaxies (refs 9, 10 and P. P. Kronberg, unpublished). After combining these data with other recent measurements[11,12] and other previously published polarisation data, we have computed rotation measures for 450 extragalactic sources. These include 108 QSOs of known redshift, of which 48 have a redshift $z > 1$—many more than the number previously in this category. Using the new larger data sample, we re-examine the RM with a particular view to reassessing the relative contributions of factors (1), (2) and (3). Of particular interest is whether any effect of an intergalactic medium (2) can be seen, since the detection of intergalactic gas by its effect on Faraday rotation of QSOs would have important cosmological significance.

Comparison of the galactic and extragalactic contributions

Figure 1 shows histograms of the observed RMs for all sources (radio galaxies, QSOs and unidentified) separately for (a) $|b| > 30°$, where the galactic contribution should be small, and for (b) $|b| \lesssim 30°$. Figure 1(a) shows that there is a clear separation into two populations; a narrow peak centred on $RM \simeq 0$ rad m^{-2} and not resolved by the bin width of 50 rad m^{-2}, and a second, broad distribution extending more or less symmetrically to approximately $\pm 1,000$ rad m^{-2}. For sources having $|b| < 30°$, the broad and symmetric group is still present, in fact the fraction of sources showing $|RM| > 200$ rad m^{-2} is virtually the same in (a) and (b)—34% and 35% respectively. There is a striking difference, however, between the distributions near $RM = 0$. What is a narrow peak of half-width ~ 45 rad m^{-2} in Fig. 1a (see inset) appears

*Visiting von Humboldt Fellow on leave from the University of Toronto.
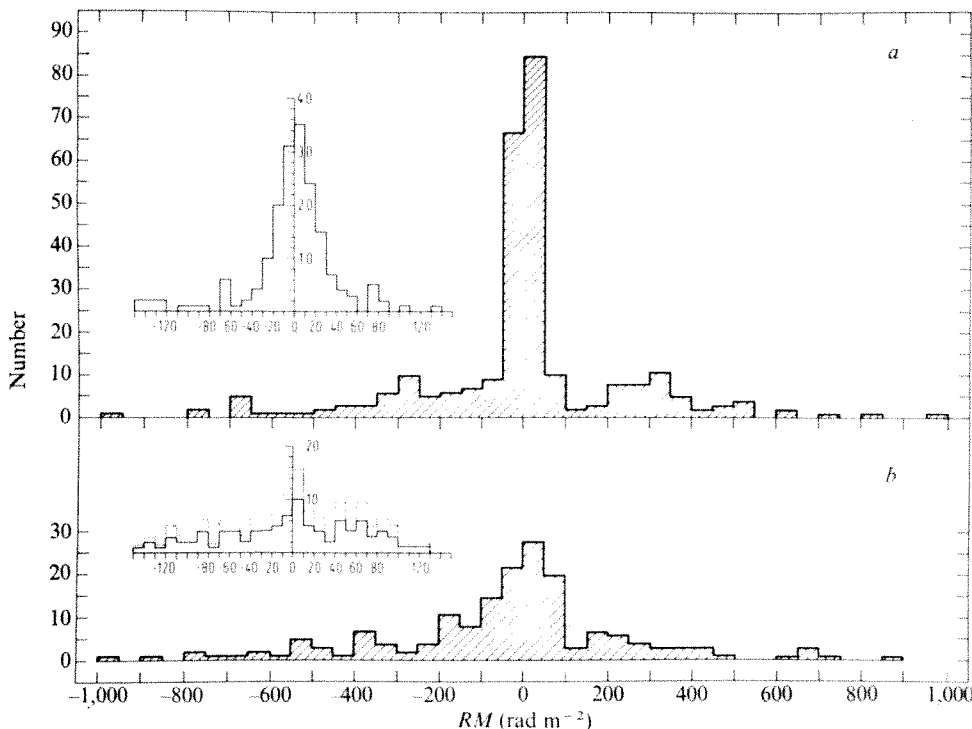
Fig. 1 Histograms showing the distribution of the rotation measures for (a) |b| > 30° (275 sources and (b) |b| < 30° (175 sources). The insets show the central peaks with intervals of 10 rad m^{-2}. The dotted profile in (b) shows the same distribution scaled by 275/175 for comparision with that in (a).

as distinctly broader distribution, with half-width ~ 120 rad m^{-2} for the low-latitude sources. Again comparing distributions (a) and (b) in Fig. 1 it is evident that sources with very low rotation measures |RM| ≲ 20 rad m^{-2} occur nearly exclusively at the higher galactic latitudes (|b| ≳ 30°), whereas those with very large rotation measures occur with equal probability everywhere in the sky. The comparison between Fig. 1(a) and (b) is revealing, and enables us to draw some new and interesting conclusions. These are: firstly, the very large (≳ 200 rad m^{-2}) values of RM are, on average, not caused by the Galaxy, since the fraction of sources with |RM| > 200 is independent of path length through the Galaxy. A two-dimensional galactic co-ordinate plot of the rotation measures (not shown) confirms that there is no large-scale order in Fig. 1b for the high RM sources, and that the large source-to-source variations of RM are just as large at high as at low galactic latitudes, and second, there is a small subset (~ 25% of the sources in our sample of 450) which have a very low intrinsic RM (≲ 10 rad m^{-2}). These form the narrow peak in Fig. 1(a) which, presumably, would be still narrower if we were to remove accurately the galactic component of RM at high latitudes.

The first conclusion is new, and contradicts the widespread belief based on earlier data[1-3] that most of the Faraday rotation was galactic. Our data show that the galactic component of RM must seldom be ≳ 60 rad m^{-2} for |b| > 10°. Because of various selection effects, there are very few RM at |b| ≲ 5°

where the galactic rotation measure may be systematically ≳ 200 rad m^{-2}. Thus, our first conclusion may not apply to sources which lie very close to the galactic plane, but this question does not affect the discussion of the extragalactic contribution. We shall cite further evidence below which suggests that the bulk of the large RM is generated in or near the sources themselves.

Having established that the Galactic rotation measure is < 200 rad m^{-2}, except perhaps at very low latitudes, we attempted to estimate the extragalactic rotation measure of all sources with known redshift—108 QSOs and 90 radio galaxies. To do this we define the galactic rotation measure (GRM) to be the average RM within a 20° diameter circle centred on the QSO, for all other sources having |RM| < 150 rad m^{-2}. By eliminating sources with |RM| > 150 rad m^{-2} we tend to eliminate sources whose RM is largely extragalactic. It is the residual rotation measure (RRM = RM−GRM) derived in this way which we analyse further below.

The distribution of *RRM* for QSOs and radio galaxies

Figure 2 shows the distributions of RRM for 118 sources having z ⩾ 0.2 for five redshift intervals. Note that the 48 sources having z > 1 are nearly equally divided among the three largest redshift intervals. If the RRM calculated for each source
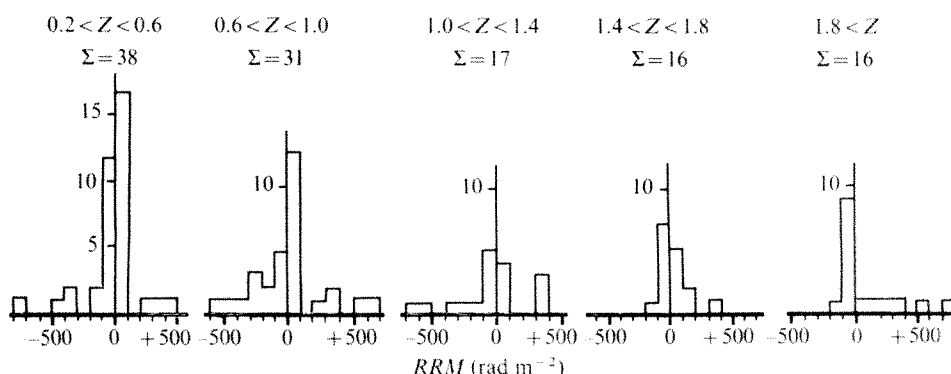
Fig. 2 Histograms showing the distribution of the residual rotation measures (RRM) for 5 equal redshift intervals beginning at z = 0.2. The formal variances are (from left to right): 4.0, 5.5, 6.6, 1.5 and 5.7 × 10^4 rad^2 m^{-4}.

were a true measure of the extragalactic component of its rotation measure, and if there were no intervening intergalactic material which contributes to the *RM*, we would attribute the observed variances solely to differences in internal Faraday rotations in the source themselves. If the internal Faraday rotation within the sources is independent of redshift (and assuming the emission line redshift also to be that of the radio emission) the variance should then decrease as $(1+z)^{-4}$. Figure 2 shows that this is clearly not the case: within the errors, the variance of *RRM* is largely independent of redshift for the 118 sources (101 QSOs) having $z > 0.2$. We also computed the *RRM* of a control sample of 90 radio galaxies with known redshift by the same method and plotted the variance over their (much smaller) range of z. There is likewise no systematic variation over the range $z \lesssim 0.4$, and the variances are also in the range $(4-7) \times 10^4$ rad^2 m^{-4} which is indistinguishably different from that for the QSOs.

We repeated the variance calculations for the same bin intervals and populations as in Fig. 2, but this time using different randomly-mixed redshifts for the sources. The variance is in the range $(5-6) \times 10^4$ rad^2 m^{-4}, and shows a slight tendency to increase at low z where there are more sources in the bin—an effect of a few anomalously large rotation measures which are more likely to fall in the low redshift bins since these have the largest populations. As a further test we recalculated the variances using 'observed' (*RM*) rather than the residual (*RRM*) rotation measures, and the average variance is only marginally greater, which is consistent with our expectation based on the results of Fig. 1.

We can conclude from this that the rotation measures have a variance $V \sim 5 \times 10^4$ rad^2 m^{-4} even after attempting to remove the galactic contribution. This noise effect is not substantially reduced if we eliminate all sources close to the galactic plane ($|b^{11}| < 30°$), where the galactic contribution is larger. The fact that the scatter is of the same order of magnitude for the (low redshift) sample of radio galaxy *RRM* as for the QSOs suggests that the large source to source variations are not a sensitive function of intergalactic path length. The overwhelming contribution to large *RM* must therefore come largely from the sources themselves. This fact was previously surmised by Reinhardt[13], and it is also consistent with the observed depolarisation rates in many extragalactic sources, which often imply a differential Faraday rotation within the sources of several hundreds of rad m^{-2}.

It is now appropriate to consider whether some other source parameter correlates with *RM* (or *RRM*), which might enable a better statistical isolation of the intergalactic contribution. If so, we could improve the sensitivity of our test for variations of $V(z)$ in Fig. 2, which are overwhelmed by the scatter from *RM* generated at or in the sources themselves. For most of the QSOs we also have measurements of $\lambda_{1/2}$, the wavelength at which the linear polarisation falls to half its maximum value. This depolarisation is most likely caused by differential Faraday rotation within the sources themselves, and if we further assume the sources to have a uniform mixture of thermal and relativistic electrons the rotation measure which is intrinsic to the source, (*IRM*), can be related to $\lambda_{1/2}$ by

$$IRM = k\lambda_{1/2}^{-2} \text{ rad m}^{-2} \tag{1}$$

where the factor k is $\simeq 1(\lambda_{1/2}$ in (m)) for most simple geometrical shapes which have an aligned internal magnetic field[14]. Obviously we do not expect the sources to behave as simple uniform mixtures of relativistic and thermal electrons in a perfectly homogeneous magnetic field, as $k = 1$ in equation (1) implies, but it is interesting to see if some statistical correlation exists. A plot of *RRM* against $\lambda_{1/2}^{-2}$ (Fig. 3) shows that there is in fact little correlation between depolarisation rate and rotation measure.

Figure 4 shows the distributions of the calculated values of *IRM* for the same redshift intervals as in Fig. 2 for the QSO sample. There is a possible tendency for the widths of the
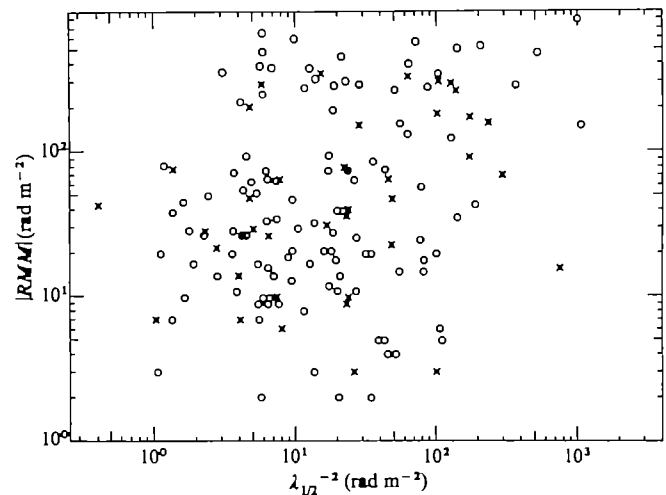
Fig. 3 A plot of *RRM* against $\lambda_{1/2}^{-2}$. $\circ \Rightarrow z < 1.0$; $\times \Rightarrow z > 1.0$.

RRM distributions in Fig. 2 to reflect those of $\lambda_{1/2}^{-2}$ in Fig. 4, but there is still too little data for us to conclude that there is a statistical dependence of V on $\lambda_{1/2}^{-2}$, particularly in view of the poor correlation in Fig. 3. If we were to conclude that the variances ($V(z)$) merely reflected the *IRM* of the sources we could nonetheless still establish an upper limit to the contribution of an intergalactic medium to the dispersion in *RRM*, since any variation of $V(z)$ from an intergalactic magneto-ionic medium would be superimposed on the effect of the varying scatter in the *IRM*. We also tested for the effect of varying populations of radio spectral index, $\alpha(Sv \propto^\alpha)$, with redshift. After splitting the QSOs according to $\alpha \lesssim -0.5$, we found no significant difference in the variances for each population, and conclude that the scatter in *RRM* in the present sample is quite insensitive to variations in radio spectral index.

These investigations show that in the absence of more and better data, a fairly large scatter in the *RM* values of QSOs with large z cannot be removed, and any contribution by Faraday-rotating intergalactic clouds would have to be $\gtrsim 100$ rad m^{-2} in the observer's frame to have a detectable effect on the formal variances. If, however, a substantial subgroup of sources with *RRM* $\simeq 0$ in fact persists out to the largest redshifts (analagous to the sharp peak near *RM* $= 0$ in Fig. 1) this would provide powerful evidence for the absence of any intergalactic gas There is some suggestion that a narrow peak near *RRM* $= 0$ does persist out to $z \simeq 2$ (Fig. 2), but our sample of highly redshifted objects is not quite large enough to provide a firm answer to this question.

Limits to the contribution from an intergalactic medium

The fact that we find no detectable increase in the variance $V(z)$ of the *RRM*s with redshift (Fig. 2) enables us to put an upper limit on any contribution caused by intervening intergalactic clouds of magnetoionic material. If the ion density of the intergalactic clouds follows the universal scale factor and the (randomly oriented) magnetic field is frozen in, then

$$n(z) \propto (1+z)^3 \tag{2}$$

$$|B|(z) \propto (1+z)^2 \tag{3}$$

and the increase in $\langle n|B_{||}|\rangle$ with redshift overwhelms the 'watering-down' effect due to the Doppler shift[15,16] ($\alpha(1+z)^{-2}$) in which case the variance should increase with redshift. For these assumptions the sensitivity of the *RM* data to an intergalactic medium increases as the source sample extends to greater redshift. For our sample, of which 48 sources have $1 < z < 2.8$, we can infer the following upper limit to an
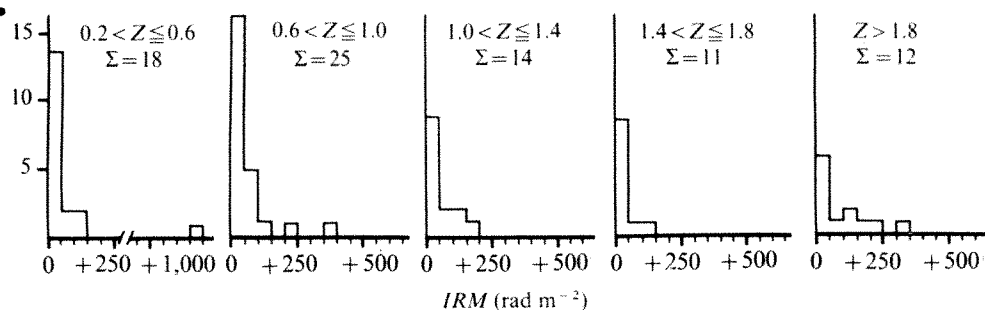
Fig. 4 Distributions of the intrinsic rotation measures of QSOs (in the observer's frame) as calculated from the assumption that $IRM = \lambda_{1/2}^{-2}$ and using the same redshift intervals as in Fig. 2. The samples here are slightly smaller than in Fig. 2, since a few sources do not have well-determined $\lambda_{1/2}$ values.

intergalactic medium in a Friedmann universe with $q_0 = 0.5$ and $H_0 = 75 \text{ km s}^{-1} \text{ Mpc}^{-1}$

$$n_0 |B_{||0}| \lesssim \frac{1.1 \times 10^{-13}}{l_0^{1/2}} \text{ gauss cm}^{-3}$$

where $n_0|B_{||0}|$ and l_0 are the mean values for the present epoch (averaged over many cells) of electron density, line of sight component of magnetic field, and cell size (Mpc) respectively. The cells are assumed contiguous. If we assume $B_{||0} = 10^{-8}$ gauss, then the corresponding density is 1.8 times the mean density

$$\left(\bar{\rho}_0 = \frac{3}{4\pi} q_0 \frac{H_0^2}{G} \right)$$

for this model of the Universe. This is not much above the mean density $\bar{\rho}_0$ and shows that with more data it should soon be possible to use rotation measure data to place useful limits on certain combinations of model universe and intergalactic medium parameters.

1 Gardner, F. F., and Davies, R. D., *Aust. J. Phys.*, 19, 129–139 (1966).
2 Gardner, F. F., Morris, D., and Whiteoak, J. B., *Aust. J. Phys.*, 22, 813–819 (1969).
3 Mitton, S., *Mon. Not. R. astr. Soc.*, 155, 373–381 (1972).
4 Vallée, J. P., and Kronberg, P. P., *Astr. Astrophys.*, 43, 233–242 (1975).
5 Mitton, S., and Reinhardt, M., *Astr. Astrophys.*, 20, 377–340 (1972).
6 Nelson, A. H., *Pub. astr. Soc. Japan*, 25, 489–494 (1973).
7 Vallée, J. P., *Nature*, 254, 23–26 (1975).
8 Kawabata, K., Fujimoto, M., Sofue, Y., and Fukui, M., *Pub. astr. Soc. Japan*, 21, 293–305 (1969).
9 Wardle, J. F. C., and Kronberg, P. P., *Astrophys. J.*, 194, 249–255 (1974).
10 Vallée, J. P., and Kronberg, P. P., *Astrophys. J.*, 193, 303–308 (1974).
11 Haves, P., Conway, R. G., and Stannard, D., *Mon. Not. R. astr. Soc.*, 169, 117–131 (1974).
12 Wright, W. E., thesis, California Institute of Technology (1973).
13 Reinhardt, M., *Astr. Astrophys.*, 19, 104–108 (1972).
14 Burn, B. J., *Mon. Not. R. astr. Soc.*, 133, 67–83 (1966).
15 Rees, M. J., and Reinhardt, M., *Astr. Astrophys.*, 19, 189–192 (1972).

Muon catalysed fusion for pellet ignition

W.P.S. Tan

Department of Natural Philosophy, University of Glasgow, Glasgow G12 8QQ, UK

We describe here a new method for achieving controlled thermonuclear fusion which appears to give a better overall energy output over input ratio than for laser driven fusion.

THE idea is to use muon catalysed fusion reactions, which can take place at a temperature lower than that required for ordinary fusion reactions, to provide the energy needed for the ignition of ordinary fusion reactions in an inertially confined deuterium–tritium pellet. The pellet is first compressed to ~ 10^3 times solid density and preheated to ~ 10^3 eV by a relatively low power laser or electron beam acting as a prepulse. This is to enhance both the catalytic and ordinary fusion reaction rates. Simultaneously, or a short time beforehand, a pulse of muons (probably > 10^{10} in < 1 ns) is injected into the pellet. The muon energy distribution is selected such that most of the muons are deposited in the core of the pellet. The distance stopped and the moderation time can be obtained as a function of the muon incident energy[1]. For example 1-MeV muons will be stopped in 10^{-3} cm in 10^{-11} s.

Basic reaction

When stopped, a muon cascades through the atomic energy levels and enters a muonic Bohr orbit about the nucleus (here a

deuteron or triton). Since the muon is 205 times heavier than the electron its Bohr orbit (mean radius 2.6×10^{-11} cm) is very close to the nucleus. The orbiting muon thus effectively screens off the Coulomb potential of the nucleus. The muon is, moreover, polarised by an incident ion, resulting in further reduction of the Gamow barrier penetration width (Fig. 1 inset) and this greatly increases the probability of the fusion reactions which can be written as

$$\left.\begin{array}{l} \mu t + d^+ \\ \mu d + t^+ \end{array}\right\} \rightarrow \alpha(3.6 \text{ MeV}) + n(14 \text{ MeV}) + \mu \qquad (1)$$
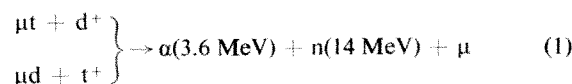
In addition to the usual 3.6-MeV α particle and 14-MeV neutron, a muon with negligible energy is also ejected. The free muon is again captured by a deuteron or triton and the cycle is repeated until it is captured by a helium double ion, in effectively another branch of the reaction (1)

$$\mu t + d^+ \rightarrow \mu He^{4+} + n \qquad (2)$$

It is then unavailable to the catalytic fusion chain unless it is stripped off in the slowing down of the helium ion.

The branching ratio between reactions (1) and (2) is 100:1,
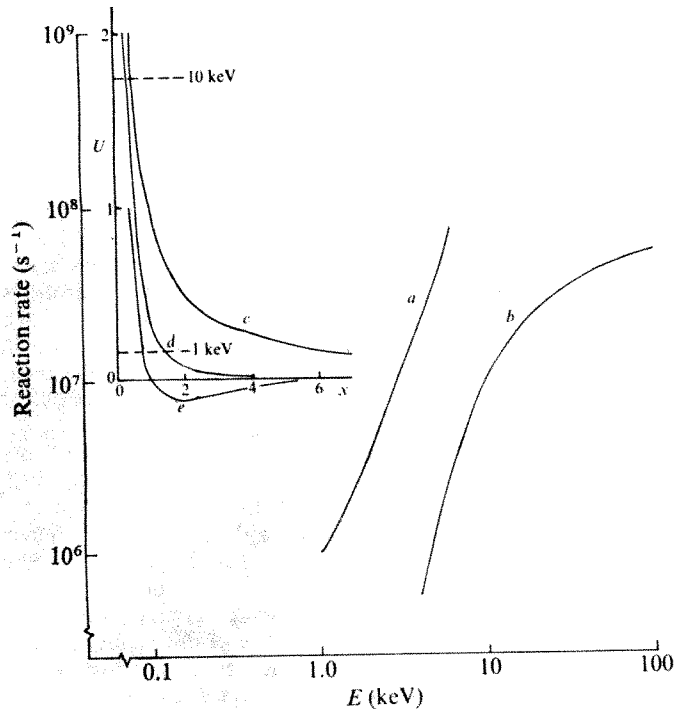
Fig. 1 Reaction rates of muon catalysed 'in flight' dt fusion reactions at solid dt density. *A*, . . . Muon catalysed reaction; *B*, . . . ordinary fusion reaction.

Inset: Potential distribution of muonic hydrogen atom and its polarisation by a positively charged nucleus. *C* . . . $U(x) = x^{-1}$; *D* . . . $U(x) = \exp(-x)/x$ (screened Coulomb potential);

E . . . $U(x) = \dfrac{1}{x} - \dfrac{1.5}{(1 + 0.127x)^4}$ (Morse potential);

$x = r/a$, $a =$ muonic Bohr radius, $U = \dfrac{E}{(e^2/a)}$

which is also the catalytic chain ratio unless the pellet core confinement time is too short for the chain to be completed. Inertia confinement times are typically $\sim 10^{-9}$ s or shorter so that the muon which has a mean half life of 2×10^{-6} s is stable in our time scales.

Mechanism

Muon catalysed fusion reactions were initially studied by Frank[2] and Alvarez *et al.*[3] and put on a firm theoretical basis by Jackson[4], Belyaev *et al.*[5], Zel'dovich and Gershtein[6] and Carter[7]. One important conclusion is that the fusion reaction (1) in the cold medium, requires first the formation of the molecular ion dμt[+]. The reaction rate is therefore controlled by the rate of formation of the molecular ion, and is therefore slow. In our scheme we propose to bypass this stage by preionising the electronic atoms.

The reaction rate, at constant density, then, is the product of two probabilities: the probability P_1 for barrier penetration and the probability P_2 of decay of the compound nucleus μdt into the right channel. $P_1(E)$, as a function of the energy, E, of the incident nucleus, has been computed by solving Schrödinger's equation to obtain the ratio of the wave amplitudes at the nucleus radius and at a large distance from it, using the Morse potential distribution, which is an approximation of the polarised screened Coulomb potential (Fig. 1 inset). For energies < 10 keV, $P_2(E)$ is not known. There is, however, a resonance 'near zero energy'[8] and it is likely that P_2 is greater than the 2×10^{-14} cm^3 s^{-1} given by Jackson[4] which is roughly valid for an energy between 10 keV and 100 keV. Assuming $P_2 = 2 \times 10^{-14}$ cm^3 s^{-1}, the reaction rates for the reaction (1) at a solid density of 5×10^{22} cm^{-3} is given in Fig. 1, together with the ordinary fusion reaction rates.

Rates

To find the number of particles undergoing fusion reactions, it is not enough to know the reaction rates, since the muonic atom may be dissociated ('ionised') in certain conditions of temperature and density. Moreover, we require the pellet to be sufficiently ionised with respect to the electrons, that is for the reactions

$$\mu + t^+ \rightleftharpoons \mu t \text{ or } \mu + d^+ \rightleftharpoons \mu d$$

$$e + t^+ \rightleftharpoons et \text{ or } e + d^+ \rightleftharpoons ed \tag{3}$$
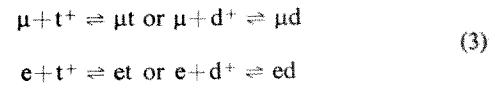
we require

$$\alpha_{\mu t} \equiv \frac{n_{\mu t}}{n_{\mu t} + n_\mu}(\text{or } \alpha_{\mu d}) \text{ and } \alpha_{t^+} \equiv \frac{n_{d^+}}{n_0}(\text{or } \alpha_{t^+}) \tag{4}$$

to be sufficiently large, where n is the number density and n_0 is the number density of dt particles in the pellet.

Now, the ionisation potentials for the μt (or μd) and et (or ed) atoms are, respectively, 2.7 keV and 13.5 eV. Applying Saha's equation on equilibrium constants to the reactions (3) we obtain $\alpha_{\mu t}$ (or $\alpha_{\mu d}$) and α_{t^+} (or α_{d^+}) (Fig. 2). The reaction $\mu + t^+ \rightleftharpoons \mu t$ (or $\mu + d^+ \rightleftharpoons \mu d$) is in equilibrium because its reaction time is $\sim 10^{-12}$ s (the case of low energy ionisation loss of ref. 1) at solid density and proportionately smaller at higher density. Thus, compared with a confinement time of 1 ns the dissociation/recombination time is small. The reaction $e + t^+ \rightleftharpoons$ et (or $e + d^+ \rightleftharpoons$ ed) has a much shorter time than 10^{-12} s at solid density since its cross section is obviously larger than for the reaction $\mu + t^+ \rightleftharpoons$ et.

Both $\alpha_{\mu t}$ (or $\alpha_{\mu t}$) and α_{t^+} (or α_{d^+}) are sufficiently large in the density and temperature range envisaged, for catalytic fusion to take place almost up to ignition point (see Fig. 2). When the catalytic reactions are nearly complete the ignition temperature would be reached and the main burn will take place by means of ordinary fusion reactions.

Output

For an idea of the overall energy output over input ratio we consider a few cases. The parametric relations used are (where number densities are for one pulse)

$$E_t = x_b N T_i/R_b = n_\mu T_r R_c x_r \tag{5}$$

Fig. 2 Ionisation fractions. $\alpha_{\mu t}$ ——— , (possible implosion path ——→——); α_{t^+} — — — — — , (possible implosion path — — → — —).
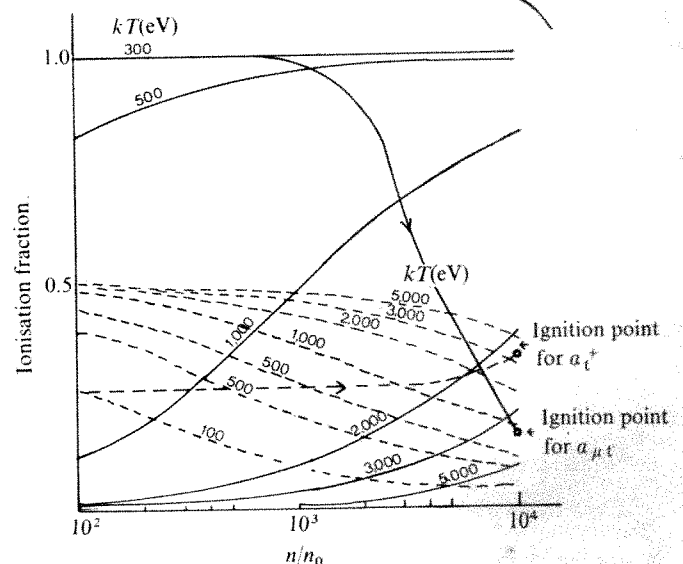
Table 1 Limits to fusion parameters

Parameter	Minimum	Maximum	Remarks
N	2×10^{19}	2×10^{22}	Lower limit comes from a pellet of 0.5 mm radius, possibly the smallest pellet consistent with focusing of muon beam
x_b	0.01	0.1	Lower limit is the 1% burn fraction as calculated by Brueckner[8] for a 0.5-mm pellet in laser driven fusion
R_b	200	2,000	Lower limit from laser driven fusion[8]
T_i	← 3.5 KeV →		$T_i = 3.5$ keV is the result of large R_b as given by Lawson[10]
T_r	3.5 MeV	17.6 MeV	Lower limit is the result of α particles stopped in core zone. Upper limit is the result of both α and n particles stopped in core zone
R_o	1	100	Upper limit has a firm theoretical basis[4,6]
x_r	0.1	1.0	Upper limit implies that all neutrons and α-particles are stopped in ignition zone (possible for 10^4 compression)
x_μ	10^{-4}	$10^{-3}(?)$	Lower limit has a firm experimental basis[9]
e_μ	10^9 eV	10^{10} eV	Upper limit is a reasonable estimate in 1957[4]

where E_t is the thermal energy required as well as the energy available for ignition. N = number of d and t particles in the pellet; x_b = fraction of these particles burned; R_b = burn–ignition particle ratio; T_i = ignition temperature in eV; n_μ = number of muons injected into the pellet and assumed also to be the number of muons collimated into a beam; T_r = energy of reaction products; R_o = catalytic chain ratio; x_r = fraction of reaction products stopped in the ignition zone.

If N_μ is the total number of muons produced from the electrical mains, e_μ is the mains energy (eV) required to produce one muon and X_μ is the fraction of muons collimated into a beam then the total input energy E_M is

$$E_M = e_\mu N_\mu = \frac{e_\mu n_\mu}{x_\mu} = \frac{e_\mu x_b N T_i}{x_\mu x_r R_b R_o T_r} \quad (6)$$

The fusion energy produced per pulse is

$$E_F \simeq e_f x_b N \quad (7)$$

Table 2 Possible parameter combination

Cases	1*	2†	1a*	2a†
N	2×10^{19}	2×10^{22}	↑	↑
x_b	0.01	0.1		
R_b	200	2,000	as in Case 1	as in Case 2
T_i(keV)	3.5	3.5		
T_r(MeV)	3.5	17.6		
R_o	100	100		
x_r	1.0	1.0	↓	↓
x_μ	10^{-3}	10^{-3}	10^{-4}	10^{-4}
e_μ	10^9	10^9	10^{10}	10^{10}
E_t(eV)	3.5×10^{19}	3.5×10^{21}	3.5×10^{19}	3.5×10^{21}
n_μ	10^{10}	2×10^{11}	10^{10}	2×10^{11}
N_μ	10^{13}	2×10^{14}	10^{14}	2×10^{15}
E_M(eV)	10^{22}	2×10^{23}	10^{24}	2×10^{24}
E_F(eV)	3.5×10^{24}	3.5×10^{28}	3.5×10^{24}	3.5×10^{28}
E_F/E_M	3.5×10^2	1.75×10^5	3.5	1.75×10^4

* Cases 1 and 1a are for the minimum size pellet.
† Cases 2 and 2a are for the maximum size pellet.
Cases 1a, 2a have worse muon collimation (x_μ) and input energy requirement (e_μ) characteristics than cases 1, 2 respectively.
E_t, n_μ, N_μ, E_M, E_F and E_F/E_M are the results of assumptions made in the values for N, x_b, R_b, T_i, T_r, R_o, x_r, x_μ and e_μ.

where $e_f = 17.6$ MeV is the energy produced in a deuterium–tritium fusion reaction, and where the contributions from the catalytic reactions are neglected since the burn–ignition ratio is normally large. The overall energy output over input ratio is thus

$$E_F/E_M = \frac{e_f}{e_\mu} \frac{x_\mu x_r R_b R_o T_r}{T_i} \quad (8)$$

The probable upper and lower limits of the parameters in equation (8) are given in Table 1. Four possible combinations of these parameters are given in Table 2. These are for two pellet sizes: a small pellet of ~ 1 mm diameter, as used in laser fusion studies, and a larger one of ~ 1 cm diameter. Though it appears from equation (8) that E_F/E_M is independent of pellet size (that is, N) in practice the parameters x_μ, x_r, R_b, R_o and T_r are probably dependent on it; x_b, R_b and T_r probably more so than the others, and are chosen accordingly in Table 2. Thus T_r is chosen to be 17.6 MeV for the larger pellet since it is thought probable that both neutrons and α particles are stopped at the core, whereas only the α particles are stopped in the case of the smaller pellet. Similarly, R_b may be an order of magnitude larger for the larger pellet if the compressed core is an order of magnitude larger than the core of the smaller pellet.

In the four case studies of Table 2, E_F/E_M varies widely from 3.5 to > 10^5 and detailed computer simulation of the implosion hydrodynamics is necessary to clarify this point, though one obvious conclusion is that an improvement in the muon collimation (that is, in x_μ) would result in a positive gain to E_F/E_M.

At this stage it may be pointed out that, following the examples of laser fusion computer studies, most of the energy absorbed by the pellet will go into the compression of the core and only a little more is needed for ignition. Thus it may be objected that there is no need for the heating contribution of muon catalysed reactions. This, however, is only true for a small pellet under conditions optimum to laser induced implosions. For a larger pellet with a large core region and with the compression ratio at the edge of the core $\ll 10^4$ the confinement time criterion may still be satisfied while substantial core heating may be necessary for ignition. This is where the muon catalysed reactions would prove useful. Also unlike laser-driven fusion, electrons which are degenerate at high densities, need not be heated at all (though they will be heated as a matter of course by the pre-pulse) and this will reduce the thermal energy requirement.

Muon production

At present muons are most efficiently produced from the decay of pions, which are themselves products of the collisions of energetic protons with a suitable target. At SINR, and elsewhere, muon 'factories' have recently come into operation, capable of producing muon pulses of > 10^7 particles from 10^{11} protons, that is, $x_\mu \simeq 10^{-4}$. The reason for such poor collimation is that as pions decay the muons are emitted isotropically. Another problem is to obtain a large pulse in a short time. Here a storage ring with a fast outlet shutter could be employed. The particles stored can be either protons or muons. The former have an infinite lifetime so that the injection current (into the ring) need only be large enough to satisfy the rate of explosions required by the reactor. On the other hand, if the muon collimation is poor, a far larger number of particles has to be stored in the ring if the particles are protons than if they are muons. Muons, when initially produced, have energies too high for the muons to be stopped by a small pellet. They will, therefore, have to be decelerated by some means, perhaps by the emission of cyclotron radiation in the storage ring. If muons are stored in a ring the injection current has a lower limit determined by the mean lifetime and the output current pulse. A muon storage ring has recently been commissioned at CERN.

1 Wightman, A. S., Phys. Rev., 71, 521–528 (1950).
2 Frank, F. C., Nature, 160, 525–527 (1947).
3 Alvarez, L. W., et al., Phys. Rev., 105, 1127–1128 (1957).
4 Jackson, J. D., Phys. Rev., 106, 330–339 (1957).

5 Belyaev, V. B., et al., Soviet Phys. JETP, 37, 1171–1178 (1960).
6 Zel'dovich, Ya.B., and Gershtein, S. S., Soviet Phys. Uspekhi, 3, 593–623 (1961).
7 Carter, B. P., Phys. Rev., 141, 863–872 (1966).
8 Brueckner, K. A., IEEE Trans. Plasma Sci., 1, 13–26 (1973).
9 Feinberg, G., and Lederman, L. M., A. Rev. nucl. Sci., 13, 431–504 (1963).
10 Lawson, J. D., Proc. phys. Soc., 70, 6–10 (1957).

Evidence for the Azores mantle plume from strontium isotope geochemistry of the Central North Atlantic

W. M. White & J.-G. Schilling

Graduate School of Oceanography, University of Rhode Island, Kingston, Rhode Island 02881

S. R. Hart

Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, DC 20015
and
Department of Earth and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Strontium isotope variations in basalts from the Mid-Atlantic Ridge and the Azores Islands support the Azores mantle plume model. The similarity of isotope ratios in Mid-Atlantic Ridge tholeiites from the Azores platform and Azores Islands alkali basalts implies that these different magma types are derived from the same mantle source but under different conditions of origin.

THE two dominant modes of igneous activity in ocean basins, oceanic island and mid-ocean ridge, produce magmas of distinctly differing geochemistry[1-3]. Mid-ocean ridge basalts are predominantly tholeiitic and have very low large ion-lithophile (LIL) element abundances and low (large ion)/(small ion) abundance ratios of LIL elements such as Rb/K, Ba/Sr, and La/Sm. On the other hand, oceanic island lavas may be of either the tholeiitic or alkaline type, both of which have distinctly higher LIL element concentrations, higher LIL element ratios, and generally higher abundances of radiogenic Sr and Pb[2]. As both MOR and oceanic island volcanics are derived from the mantle, this contrast implies significant mantle heterogeneity and differences in mantle evolution.

The Azores archipelago consists of nine volcanic islands which emerge from a shallow submarine plateau straddling the Mid-Atlantic Ridge (MAR) and the triple junction of the North American, Eurasian, and African plates[4] at approximately 40°N. Because of the closely related setting of the Azores Islands, the submerged plateau, and the transecting MAR axis, the region is excellent for studying the geochemical contrasts and spatial relations of mid-ocean ridge and oceanic island basalts.

A previous study by Schilling[5] showed that tholeiitic basalts erupted along the MAR transect of the Azores platform have patterns enriched in light rare earths which grade regularly southwards along the ridge axis to patterns depleted in light rare earths, typical of mid-ocean ridge basalts, at 33°N. Because these trace element variations could not reasonably be accounted for by different degrees of partial melting or fractional crystallisation, Schilling interpreted this gradient as resulting from a mantle plume[7] enriched in LIL elements, rising beneath the Azores and mixing southwards along the ridge with the source of normal ridge basalts 'depleted' in LIL elements in the low velocity layer. (La/Sm)$_{e.f.}$ ratios were also found to decrease northwards from 45°N to the Gibbs Fracture Zone[6].

Since trace element ratios can be affected to some extent by processes such as fractional crystallisation and partial melting, further documentation of the geochemical variation observed

by Schilling was desirable. In particular, whereas the La/Sm ratio in tholeiites may be taken as representative of the source (if they are produced by at least 20% melting of a peridotitic mantle), this is not the case for magmas which may be produced by smaller degrees of melting, such as alkali basalts[8]. For this reason, Schilling did not study the relationship between the alkaline Azores island lavas and lavas from the MAR transect of the Azores platform. Sr-isotope ratios, because of their relative insensitivity to magmatic processes, are useful in examining these relationships.

Fig. 1 Map of the study area. ●, RV Trident dredge stations; ○, other dredge stations; ×, locations of dredge stations for which $^{87}Sr/^{86}Sr$ was reported by Hart et al.[12].
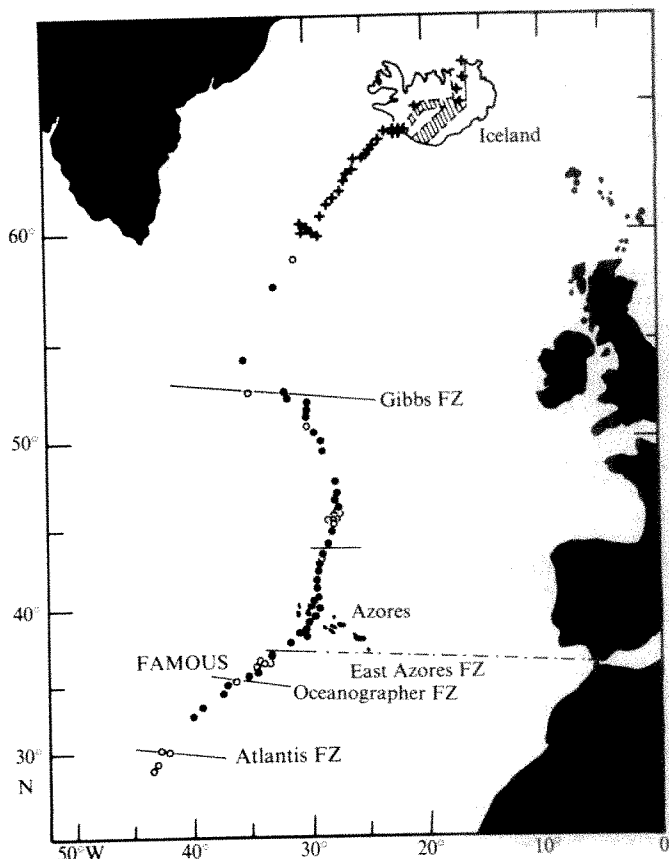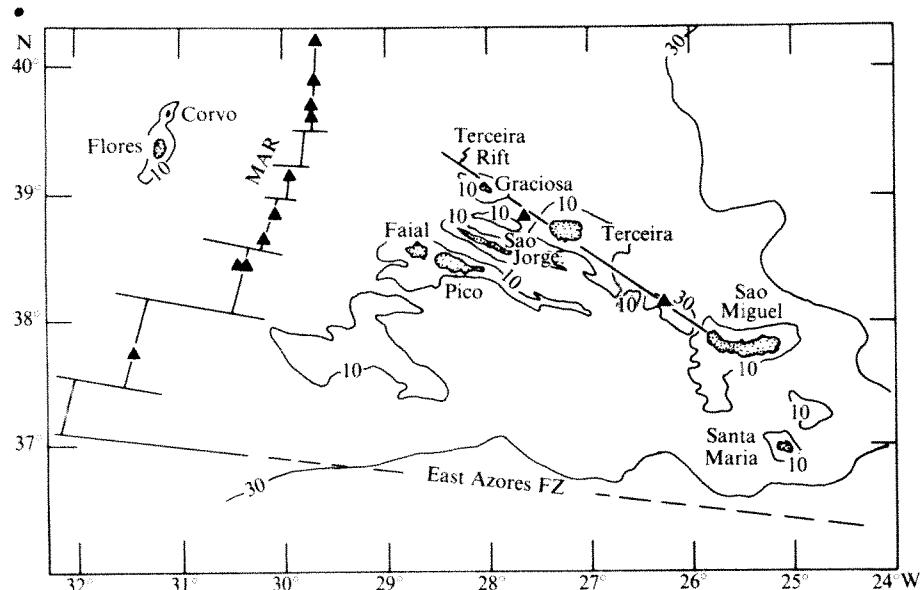
Fig. 2 Map of the Azores region. ▲, Dredge stations. 1,000- and 3,000-m contours are shown.

We now report the results of a Sr-isotope study of lavas erupted along the MAR from 29°N to 60°N, on the Azores Islands, and in the Terceira trough. The analyses were carried out on powder splits of the same MAR dredge samples on which rare earths had been determined previously by Schilling[5] and Schilling *et al.*[6]. The present study further documents geochemical variations in the region and provides a test of the Azores mantle plume hypothesis and the mixing model[5] based on rare earth elements. We also examine the relationship of alkali lavas erupted on the Azores Islands to the tholeiitic basalts erupted on the nearby MAR.

Seventy-nine samples from the rift valley of the Mid-Atlantic Ridge and the Terceira trough were analysed. Locations of the dredge stations are shown in Figs 1 and 2. Samples were dredged by RV Trident of the University of Rhode Island during cruises TR89, TR100, TR119, TR122, TR123, and TR154. Additional samples from other dredge stations were provided by other investigators. Thirty-five samples from the Azores Islands were also analysed—their locations are shown in Fig. 3. Practically all the samples from the MAR are tholeiitic basalts, but rare transitional and one mildly alkaline basalt occur near 35°N (just north of the Oceanographer Fracture Zone) and at 45°N[5,9,10]. Lavas from the Azores Islands are alkalic.

[87]Sr/[86]Sr ratios were measured using the 6-inch mass spectrometer at the Department of Terrestrial Magnetism, Carnegie Institution of Washington. The analytical technique has been described elsewhere[11,12]. Our data from Terceira Island and the southern Reykjanes Ridge agree with values reported by O'Nions and Pankhurst[13], after correction for an apparent interlaboratory bias of 0.014%. Because of the low Rb/Sr ratios and the relatively young age of the basalts, isotope ratios in these samples can be considered 'initial'. The comenditic lavas have high Rb/Sr ratios, however, and the [87]Sr/[86]Sr ratios reported for these samples cannot be strictly considered 'initial'.

Geochemical variation along the ridge

The [87]Sr/[86]Sr profile of basalts erupted along the ridge is shown in Fig. 4. Two distinct maxima are apparent, one at 44–45°N, the other over the Azores platform (39°N). The maximum over the Azores platform seems slightly higher (about 0.70345) than that at 45°N (about 0.70340). Transition zones exist north of 45°N and south of 39°N with the [87]Sr/[86]Sr ratios decreasing to typical ridge values. Ten analysed samples from the FAMOUS region, situated in the geochemical transition zone south of the Azores, fall in a narrow range of intermediate values in spite of a large variation in bulk chemistry[14]. The profile of (La/Sm)[e.f.] (ref. 5) along this section of the ridge, as

well as those of alkali and alkaline earth trace elements and ratios such as Rb/K, Ba/K, and Rb/Sr are similar to the strontium isotope profile[15]. All these different sets of data suggest that the areas south of 33°N and between 48°N and 60°N are 'normal ridge segments'.

[87]Sr/[86]Sr ratios for samples from three dredge stations in the 35°N–Oceanographer Fracture Zone area are anomalously high. Although these samples show some evidence of secondary alteration, their K/Cs ratios (a good indicator of alteration[16]) are relatively high and suggest that these high [87]Sr/[86]Sr ratios are not due to secondary alteration. Nor do the anomalously high strontium isotope ratios in this area represent a difference between rift valley and fracture zone volcanism. One dredge station is located within the MAR rift valley just north of the fracture zone. The second is located south of the fracture zone, though not necessarily in the rift (which could not be identified because of the complex topography of the area). The third is located within the fracture zone[17].

[87]Sr/[86]Sr values for station TR123 4D (33°N) are also anomalously high. These samples are quite unusual in composition, having extremely low LIL element concentrations and low ratios such as La/Sm and Rb/K. Petrologically, they are Al-rich picrites similar to rocks from DSDP sites 3-14 and 3-18[8,18] and contain Al-rich spinels[19].

Variation patterns in the Azores

The samples from the Azores Islands vary from alkali basalts to highly differentiated comendites. Submarine basalts from the Terceira trough are tholeiitic[20]. Initial [87]Sr/[86]Sr ratios from six of the nine islands fall in the narrow range of 0.70332 to 0.70354, and thus are identical to [87]Sr/[86]Sr ratios of basalts from the MAR transect of the platform (Fig. 5). Figure 5 shows also that there is no systematic change in [87]Sr/[86]Sr ratios longitudinally away from the ridge. Both island lavas and tholeiites from the ridge transect of the platform have strontium isotope ratios which are distinct from normal ridge segments.

Flores and Corvo are the only two islands west of the MAR. [87]Sr/[86]Sr values for these two islands range from 0.70332 to 0.70352, with no apparent isotopic differences between the two islands. [87]Sr/[86]Sr ratios from Faial and Pico range from 0.70347 to 0.70394, with values from Faial tending to be somewhat higher than those from Pico. Sao Jorge has [87]Sr/[86]Sr ratios which range from 0.70332 to 0.70354.

Graciosa, Terceira and Sao Miguel lie along the Terceira trough, which may be a secondary spreading centre[4]. Lavas from Graciosa and Terceira, and submarine tholeiites from the trough have [87]Sr/[86]Sr ratios in the range 0.70336 to 0.70354, with the exception of two comendites from Graciosa and
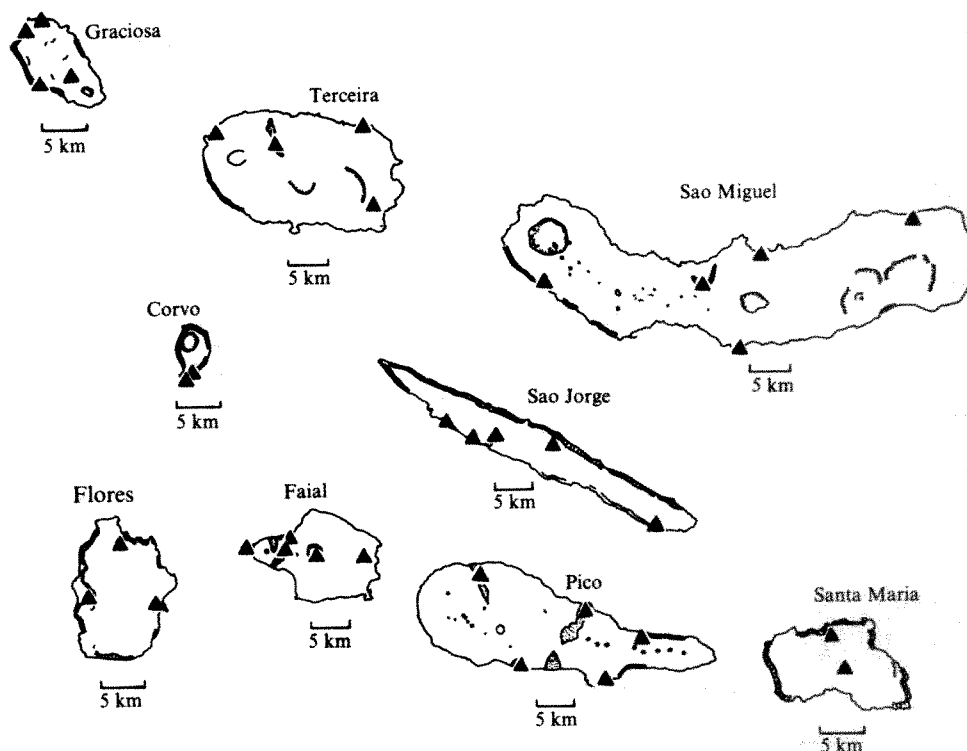
Fig. 3 Sample location map for the Azores Islands. ▲, Sample locations; dotted areas, historical lava flows.

Terceira. The high Rb/Sr ratios, 8 and 10, in the two comendites suggest that their high $^{87}Sr/^{86}Sr$ ratios may be the result of ^{87}Rb decay since differentiation of the magma. If the initial ratios of these two samples were near 0.70350, then the apparent ages are 2.5 Myr for the Graciosa sample and 0.75 Myr for the Terceira sample. Field relationships suggest that these ages may be reasonable.

Sao Miguel, which is apparently located at the intersection of the Terceira, Sao Jorge and Faial–Pico tectonic lineaments[21], is geochemically quite distinct from the other islands. One basalt has an $^{87}Sr/^{86}Sr$ ratio of 0.70337, whereas four other values range from 0.70434 to 0.70525. Samples from Santa Maria have $^{87}Sr/^{86}Sr$ ratios of 0.70336 and 0.70352.

Thus Faial, Pico and, particularly, Sao Miguel are the only islands where initial $^{87}Sr/^{86}Sr$ ratios are markedly dissimilar to those of the adjacent MAR. All these samples are fresh and the high $^{87}Sr/^{86}Sr$ ratios cannot be explained by post-eruptional alteration.

Comparison with Iceland

The maximal $^{87}Sr/^{86}Sr$ and $(La/Sm)_{e.f.}$ values for the ridge transect of the Azores platform (0.70345 and 2.8 (ref. 5) respectively) are considerably higher than values for the ridge transect across Iceland (0.70305 (ref. 12) and 1.3 (ref. 22)

respectively). This is readily apparent for $^{87}Sr/^{86}Sr$ in Fig. 4. Additionally, although the Iceland–Reykjanes Ridge transition zone between the 'undepleted' Icelandic basalts and the 'depleted' ridge basalts extends only some 400 km, the transition zone between the Azores platform and the 'normal' ridge to the south extends some 1,000 km. Thus the geochemical anomaly beneath the Azores is both richer in LIL elements and larger than the anomaly beneath Iceland.

Single or multiple mantle source?

The large scale geochemical gradients along the Reykjanes Ridge reported by Schilling[22], Hart *et al.*[12] and Sun *et al.*[23] have been explained by these authors in terms of mixing of two mantle sources. This model was challenged by O'Hara[24, 25] who suggested that the LIL element variations along the Reykjanes Ridge could be produced by fractional crystallisation. O'Hara[24] dismissed the isotope evidence, stating that the "isotope ratios had changed during the manifest fractional crystallisation." As O'Hara has not yet proposed a mechanism capable of producing such changes and as we know of no evidence to support this contention, we consider it improbable that Sr isotope ratios can be affected by fractional crystallisation.

Other alternatives to the multiple-source interpretation of the data from the Reykjanes Ridge have been based on possible

Fig. 4 Variation of $^{87}Sr/^{86}Sr$ with latitude along the Mid-Atlantic Ridge. Precision is indicated by the error bar (upper left) and is ±0.00006 (±2 s.d.). Measured values for the E and A NBS-SRM 987 standards were 0.70792 ±2 (six analyses) and 0.71017 ±5 (six analyses) (2 s.d. of the mean). All values were normalised to 86/88 = 0.1194 and are reported relative to a value of 0.70800 for the E and A standard. Data previously reported by Hart *et al.*[12] are shown for comparison (+). The dashed line is interpretive.
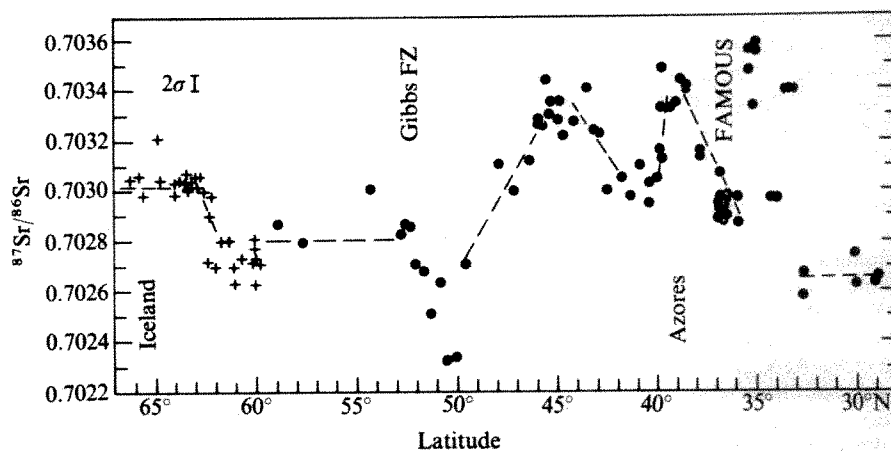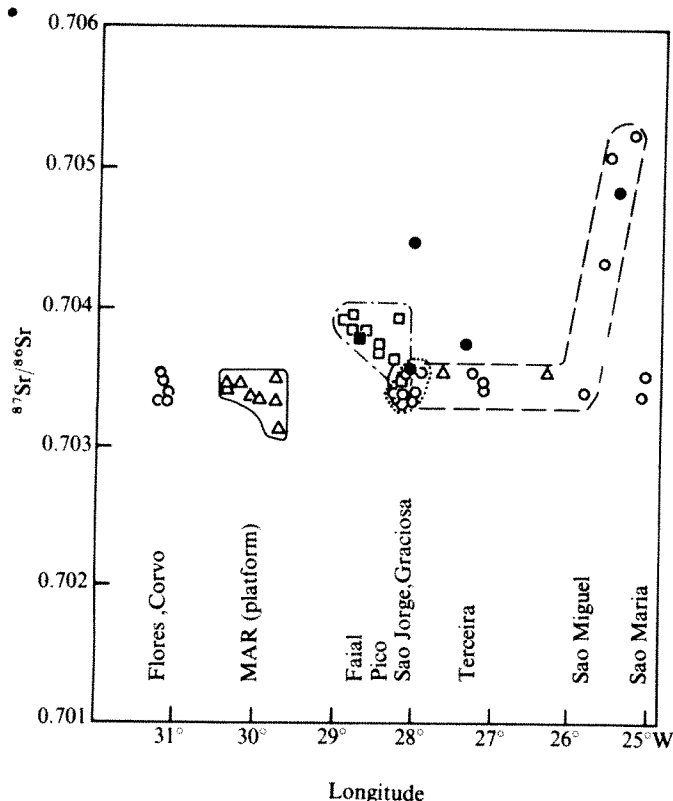
Fig. 5 Variation of $^{87}Sr/^{86}Sr$ with longitude across the Azores plateau. $\triangle$, Dredge stations; $\square$, Faial and Pico; hexagons, Sao Jorge; $\bigcirc$, Flores, Corvo, Santa Maria, Graciosa, Terceira and Sao Miguel. Open symbols are basalts and trachybasalts, filled symbols are trachytes. Lines enclose samples from the principal tectonic features of the Azores[21]: - - -, Terceira Rift; $\cdots$, Sao Jorge; $\cdots\cdots$, Faial–Pico; ——, MAR.

disequilibrium melting of a phlogopite-bearing mantle[13,26-28]. Disequilibrium melting models such as these are not supported by our results in the Azores region. The correlation of $^{87}Sr/^{86}Sr$ with elevation along the Reykjanes Ridge, on which Flower *et al.*[26] based their model, is poor or nonexistent in the present results. The minimum depth of the rift valley of the MAR transect of the Azores platform is about 1,100 m, yet the strontium isotope ratios of these samples are considerably higher than those of subaereal Icelandic basalts. Sigvaldason's model[27], suggesting that the Iceland mantle plume feeds volcanism in all of the North Atlantic, cannot explain the high $^{87}Sr/^{86}Sr$ values at 45°N and the Azores platform, or the associated gradients.

O'Nions and Pankhurst[13] have developed a disequilibrium melting model for a phlogopite-bearing mantle. A corollary of this model requires that the $^{87}Sr/^{86}Sr$ ratio of the melt should decrease with increasing degree of melting. If alkali basalts are produced by smaller degrees of melting than tholeiites, as suggested by experimental data[29], alkali basalts should have higher strontium isotope ratios than tholeiites derived from the same source. Contrary to this prediction, however, most alkali lavas from the Azores Islands have $^{87}Sr/^{86}Sr$ ratios nearly identical to those of tholeiites from the MAR transect of the Azores plateau. In addition, basalts of tholeiitic composition dredged from the Terceira trough have isotope ratios within the range of alkali basalts from the nearby islands of Terceira and Graciosa.

A number of other serious objections to the disequilibrium melting hypothesis have been raised by Hofmann and Hart[30]. In view of these objections and the above discussion, the disequilibrium melting model seems inapplicable to our results. We conclude that the observed variation in lavas from the

MAR and the Azores Islands does indeed reflect geochemical variation in the mantle.

Any model which attempts to account for the isotopic variation along the ridge must also account for (1) the increases in trace elements and trace element ratios towards the Azores[5,15]; (2) petrological and petrochemical variation along the ridge (ref. 19 and J.-G. S. and W. M. W., to be published); (3) the broadening and shoaling of the MAR in this region; and (4) the thicker crust present beneath the Azores platform by comparison with typical oceanic crust[31].

At present, the model we prefer to account for these observations is that of a mantle plume[5,7] rising beneath the Azores (with a possible separate plume at 45°N) and mixing with the source (depleted in LIL elements) of normal ridge basalts. Reasons for the choice of this model over other models such as a vertically or horizontally heterogeneous static mantle have been discussed elsewhere[3,5,12,15,22]. We interpret the contrast of the alkalic nature of the Azores Islands lavas with the tholeiitic character of basalts from the MAR transect of the platform as being due primarily to differences in conditions of magma generation and evolution, such as depth and extent of melting and subsequent crystal fractionation, rather than differences in mantle source chemistry[29]. These different conditions of magma formation generally produce some enrichment in LIL elements in the melts but do not produce isotopic differences. Higher Sr isotope ratios on Faial, Pico and, particularly, Sao Miguel imply considerable local mantle heterogeneity in this region, however.

Strontium isotope ratios do not decrease longitudinally across the platform (a distance of over 500 km), whereas they begin decreasing along the MAR within 100 to 200 km from the centre of the platform. This asymmetry may result from the fact that although there is mixing of the plume and depleted low velocity layer sources along the ridge, only the plume contributes material to island volcanism. Our analyses of island lavas include both historical flows and flows which are from some of the oldest volcanic sequences of the archipelago (Santa Maria and north-eastern Sao Miguel[32]) and reveal no particular systematic change with time. Thus the plume must have been present beneath the Azores for at least the past 4–6 Myr, and probably longer, judging from rare earth data in DSDP Leg 37 basalts[33].

1 Engel, A. E. J., Engel, C. G., and Havens, R. G., *Bull. geol. Soc. Am.*, 76, 719–734 (1965).
2 Gast, P. W., *Geochim. cosmochim. Acta*, 32, 1057–1086 (1968).
3 Hart, S. R., *Yb. Carnegie Instn Wash.*, 70, 353–355 (1971).
4 Krause, D. C., and Watkins, N. D., *Geophys. J. R. astr. Soc.*, 19, 261–283 (1970).
5 Schilling, J.-G., *Earth planet. Sci. Lett.*, 25, 103–115 (1975).
6 Schilling, J.-G., Johnson, D. G., and Johnston, T. H., *Trans. Am. geophys. Un.*, 55, 294 (1974).
7 Morgan, J. W., *Nature*, 230, 42–43 (1971).
8 Schilling, J.-G., *J. geophys. Res.*, 80, 1459–1473 (1975).
9 Aumento, F., *Can. J. Earth Sci.*, 5, 1–21 (1968).
10 Muir, I. D., and Tilley, C. E., *J. Petrol.*, 5, 409–434 (1964).
11 Hart, S. R., and Brooks, C., *Geochim. cosmochim. Acta*, 38, 1799–1806 (1974).
12 Hart, S. R., Schilling, J.-G., and Powell, J. L., *Nature phys. Sci.*, 246, 104–107 (1973).
13 O'Nions, R. K., and Pankhurst, R. J., *J. Petrol.*, 15, 603–634 (1974).
14 White, W. M., and Bryan, W. B., *Bull. geol. Soc. Am.* (in the press).
15 White, W. M., Hart, S. R., and Schilling, J.-G., *Yb. Carnegie Instn Wash.*, 74, 224–234 (1975).
16 Hart, S. R., *Yb. Carnegie Instn Wash.*, 68, 403–408 (1970).
17 Shibata, T., and Fox, P. J., *Earth planet. Sci. Lett.*, 27, 62–72 (1975).
18 Frey, F. A., Bryan, W. B., and Thompson, G., *J. geophys. Res.*, 79, 5507–5527 (1974).
19 Sigurdsson, H., and Schilling, J.-G., *Earth planet. Sci. Lett.*, 29, 7–20 (1976).
20 Tapia, M. D. M., thesis, Univ. Rhode Island (1976).
21 Agostinho, J., *Bull. Volcanol.*, 8, 123–138 (1936).
22 Schilling, J.-G., *Nature*, 242, 565–571 (1973).
23 Sun, S. S., Tatsumoto, M., and Schilling, J.-G., *Science*, 190, 143–147 (1975).
24 O'Hara, M. J., *Nature*, 243, 507–508 (1973).
25 O'Hara, M. J., *Nature*, 253, 708–710 (1975).
26 Flower, M. J. F., Schmincke, H.-U., and Thompson, R. N., *Nature*, 254, 404–406 (1975).

27 Sigvaldason, G. E., *J. Petrol.*, **15**, 497–524 (1974).
28 Sigvaldason, G. E., Steinthorasson, S., Oskarsson, N., and Imsland, P., *Nature*, **251**, 579–582 (1974).
29 Green, D. H., and Ringwood, A. E., *Contr. Miner. Petrol.*, **15**, 103–190 (1967).
30 Hofmann, A. W., and Hart, S. R., *Yb. Carnegie Instn Wash.*, **74**, 195–210 (1975).

31 Searle, R. C., *Geophys. J. R. astr. Soc.*, **44**, 537–546 (1976).
32 Abdel-Monem, A. A., Fernandez, L. A., and Boone, G. A., *Lithos*, **8**, 247–254 (1975).
33 Schilling, J.-G., Kingsley, R., and Bergeron, M., in *Init. Rep. Deep Sea Drilling Project*, 34 (edit. by Melson, W. G., *et al.*) (in the press).

An enzyme isolated from arteries transforms prostaglandin endoperoxides to an unstable substance that inhibits platelet aggregation

S. Moncada, R. Gryglewski, S. Bunting & J. R. Vane

Wellcome Research Laboratories, Langley Court, Beckenham, Kent BR3 3BS, UK

Microsomes prepared from rabbit or pig aortas transformed endoperoxides (PGG_2 or PGH_2) to an unstable substance (PGX) that inhibited human platelet aggregation. PGX was 30 times more potent in this respect than prostaglandin E_1. PGX contracted some gastrointestinal smooth muscle and relaxed certain isolated blood vessels. Prostaglandin endoperoxides cause platelet aggregation possibly through the generation by platelets of thromboxane A_2. Generation of PGX by vessel walls could be the biochemical mechanism underlying their unique ability to resist platelet adhesion. A balance between formation of anti- and pro-aggregatory substances by enzymes could also contribute to the maintenance of the integrity of vascular endothelium and explain the mechanism of formation of intra-arterial thrombi in certain physiopathological conditions.

THE prostaglandin (PG) endoperoxides (PGG_2 and PGH_2) are generated from arachidonic acid by the membrane-bound cyclo-oxygenase enzyme[1-3], and subsequently transformed to $PGF_{2\alpha}$, E_2 or D_2 (or to their 15-hydroperoxy or 15-keto derivatives) or to 12-L-hydroxy-5,8,10-heptadecatrienoic acid (HHT) and malondialdehyde[1,3,4]. Recently a novel transformation of PGG_2 and H_2 to a non-prostaglandin compound "thromboxane A_2" (TXA_2) has been reported[3,5], and a microsomal thromboxane synthetase system in blood platelets has been identified and characterised[6,7]. Most of the activity associated with "rabbit aorta contracting substance" or RCS[8,9] is now thought to be due to TXA_2 (ref. 10). TXA_2 shares with prostaglandin endoperoxides two important biological properties; they both contract strips of rabbit aorta[6,7,10,11], and cause platelet aggregation *in vitro*[5,12].

We have discovered that blood vessel microsomes contain an enzyme that transforms PG endoperoxides to an unstable principle which relaxes some blood vessels and prevents platelet aggregation.

Preparation of enzyme and bioassay

Pig or rabbit aortas (one experiment) were stripped of adventitia, snap-frozen in liquid nitrogen, crushed into a fine powder, resuspended (1:4, w/v) in 0.05 M Tris buffer (*p*H 7.5) and homogenised at high speed in a Polytron homogeniser. The homogenate was centrifuged at 1,000*g* for 15 min and the resulting supernatant centrifuged again at 10,000*g* for 5 min. The 10,000*g* pellet was also discarded and after further centrifugation at 100,000*g* for 60 min was resuspended in deionised water and lyophilised. Yield averaged 150 mg powder (51% protein)[13] per 100 g aortic tissue. Electron microscopic examination showed that this fraction was composed mainly of microsomes although there was a small contamination with mitochondria.

Activity of PGs and products of the PG endoperoxides was assayed using the cascade superfusion technique[14]. Tissues were selected from spirally cut strips of rabbit aorta, pulmonary artery, mesenteric artery, coeliac artery and vena cava, as well as rat stomach strip, rat colon, chick rectum, guinea pig ileum and guinea pig tracheal chain. These were superfused in cascade with Krebs' solution (at 37 °C) at 10 ml min⁻¹, containing a mixture of antagonists[15] and indomethacin (1 µg ml⁻¹). Changes in length of the tissues were detected using Harvard smooth muscle transducers and recorded at an overall magnification of 1–4 times on a Watanabe six-channel pen recorder.

The most useful selection of assay tissues was rabbit aorta, rat colon and rat stomach. Rabbit aorta is contracted by PG endoperoxides and TXA_2 but not by PGE_2, $PGF_{2\alpha}$ or PGD_2. Rat colon is contracted by PGE_2 and $PGF_{2\alpha}$ but not by PGH_2, PGG_2 or TXA_2 (refs 9 and 16). Rat stomach strip is contracted to different degrees by all of the above substances.

Aortic microsomes (10–1,000 µg) were incubated with substrate in 100 µl 0.05 M Tris buffer (*p*H 7.5). No cofactors were added. The following substances were tested as substrates: arachidonic acid (1–10 µg ml⁻¹), prostaglandins E_2, $F_{2\alpha}$, D_2, G_2 and H_2 (0.1–1.0 µg ml⁻¹). Reaction sets were incubated

Fig. 1 Bioassay using rabbit aortic strip (*a*) and rat colon (*b*). PGE_2 contracts rat colon whereas PGG_2 contracts rabbit aorta. PGG_2 (0.5 µg) was incubated in 0.5 ml Tris buffer (0.5 M, *p*H 7.5) at 22 °C, and 100-µl aliquots were tested on the tissues at 5, 10, 20 and 30 min. The spontaneous disappearance of PGG_2 (decrease in rabbit aorta contraction) was associated with an appearance of PGE- and F-like activity (increase in rat colon contraction).

Incubation time (min)
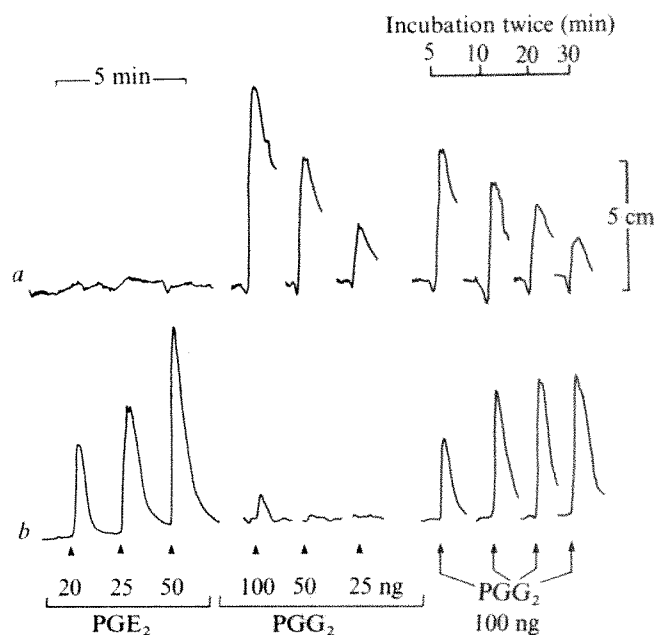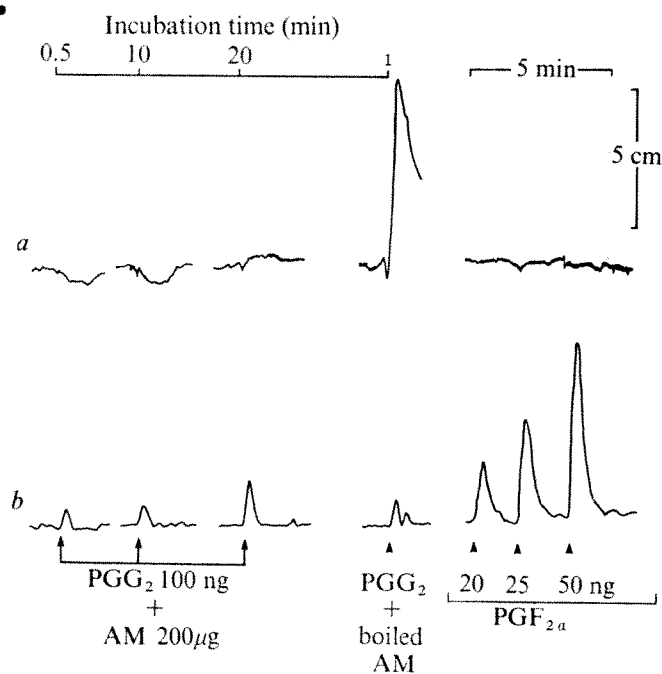
Fig. 2 Effect of aortic microsomes (AM) on PGG_2. PGG_2 (0.5 µg) was incubated with AM (1 mg) in 0.5 ml of Tris buffer (0.05 M, pH 7.5) at 22 °C and 100-µl aliquots were tested for activity on the tissues at 0.5, 10 and 20 min. *a*, Rabbit aortic strip; *b*, rat colon. In the presence of AM the contractile activity of PGG_2 on rabbit aortic strip disappeared within 0.5 min but no PGE- or F-like activity was formed even when incubated for 20 min. Boiled AM (200 µg) incubated with PGG_2 (100 ng) in 100 µl Tris buffer at 22 °C for 1 min did not affect activity of PGG_2 on rabbit aorta. The last section shows the selective effect of $PGF_{2\alpha}$ on the rat colon.

for 0.2–30 min at room temperature (20–22 °C). Aliquots of the incubation mixture were tested directly on the assay tissues to detect changes in activity induced by the enzyme. PGG_2 and PGH_2 are unstable in aqueous solution, with a half life of 7–8 min at 22 °C (ref. 7). As they decompose (Fig. 1), so the activity on rabbit aorta decreases and PGE- and F-like activity on rat colon increases.

When incubated with aortic microsomes (3–4 mg ml^{-1}) for 0.25–0.5 min at 22 °C, the contractor activity of PGG_2 or PGH_2 (1 µg ml^{-1}) disappeared, without any increase in PGE- or F-like activity on the rat colon (Fig. 2). There was also no increase in malondialdehyde formation, as estimated by the thiobarbituric acid method[17]. The endoperoxides were transformed, however, into a substance with a different spectrum of biological activity. For convenience, we shall refer to this unknown substance as PGX.

PGE_2, $PGF_{2\alpha}$, PGD_2 (0.5–5 µg ml^{-1}) or arachidonic acid (1–25 µg ml^{-1}) were incubated with aortic microsomes for up to 30 min at 22 °C; since there was no detectable change in biological activity, it was concluded that these substances were not substrates for the enzyme which generates PGX.

The enzymic nature of the transformation by aortic microsomes of PGG_2 and PGH_2 to PGX was substantiated by the following observations: boiled microsomal preparations had no activity, and non-enzymic proteins (for example, bovine serum albumin) did not affect the spontaneous decay of PGH_2 during 1 min of incubation.

Effects of PGX

PGX (100 ng), whether extracted (see below) or not, did not contract strips of rabbit aorta, pulmonary artery or vena cava. PGX relaxed strips of rabbit mesenteric and coeliac arteries. PGX contracted rat stomach strip, chick rectum, guinea pig tracheal chain and guinea pig ileum, although its potency was less than that of PGH_2 or PGG_2. Rat colon was not contracted by endoperoxides or by PGX (100 ng), except after

spontaneous decomposition of the prostaglandin endoperoxides (Fig. 1), when there was a 25–40% conversion to PGE- or F-like substances.

Aggregation of platelets in 1 ml fresh human platelet plasma (PRP) was monitored in a Born aggregometer[18]. Platelet aggregation was induced by arachidonic acid (100–600 µg or 0.3–1.8 mM). Aortic microsomes (5 mg) were incubated with 1 µg endoperoxide (PGG_2 or PGH_2) in 1 ml 0.05 Tris buffer (pH 7.5) for 2 min at 22 °C. Doses of PGX stated below assume complete conversion of the endoperoxide. An aliquot containing 0.25–10 ng PGX was added to the PRP 0.5–1 min before the aggregatory agent. In another set of experiments the products formed during incubation of aortic microsomes with PGG_2 or PGH_2 (incubation time 1–2 min at 22 °C) were rapidly extracted with ice-cold diethyl ether. After evaporation of the ether the residue was either dissolved in ice-cold Tris buffer and immediately used in platelet aggregation studies or was dissolved in anhydrous acetone and stored at −20 °C for future use. These extracts were added to PRP 1 min before the aggregatory agent.

An immediate anti-aggregatory effect of the fresh reaction mixture of aortic microsomes with PGH_2 or PGG_2 was observed. The lowest anti-aggregatory concentration was obtained at concentrations of 0.5–5 ng PGX ml^{-1}. PGX was about 30 times more potent than PGE_1 and 5–20 times more potent than PGD_2 as an anti-aggregatory agent (Fig. 3). PGX was unstable and its anti-aggregatory activity disappeared after standing for 20 min at 22 °C or after boiling for 0.2 min. Aortic microsomes (50 µg ml^{-1}) alone could induce aggregation in some PRP. The products of spontaneous degradation of PGG_2 (100 ng ml^{-1}) had no anti-aggregatory activity.

Diethyl ether extracts of PGX also inhibited platelet aggregation induced by arachidonic acid. The effective anti-aggregatory concentrations of the extracted PGX ranged from 1–10 ng ml^{-1} (assuming complete extraction into ether and complete dissolution in Tris buffer).

The anti-aggregatory activity of the extracted PGX disappeared on boiling (15 s) or on standing at 22 °C for 20 min, but did not deteriorate when the buffered solution was kept on ice for 2 h or when the substance was dissolved in dry acetone and stored at −20 °C for several days.

PGX is different from the other products of PG endoperoxides so far described. Its biological properties on the isolated tissues, its instability and its potent anti-aggregatory activity distinguish it from PGE_2, $PGF_{2\alpha}$, TXA_2 or TXB_2. PGE_2,

Fig. 3 Transformation of PGG_2 activity by different concentrations of aortic microsomes, incubated with 100 ng of PGG_2 in 100 µl Tris buffer for 1 min at 22 °C, and bioassayed using rabbit aortic strip. The reduction in contractile activity of PGG_2 was calculated from the standard dose–response curve for PGG_2 as the percentage transformation of the endoperoxide. Each point on the graph represents the mean ±s.e. of five–seven experiments.
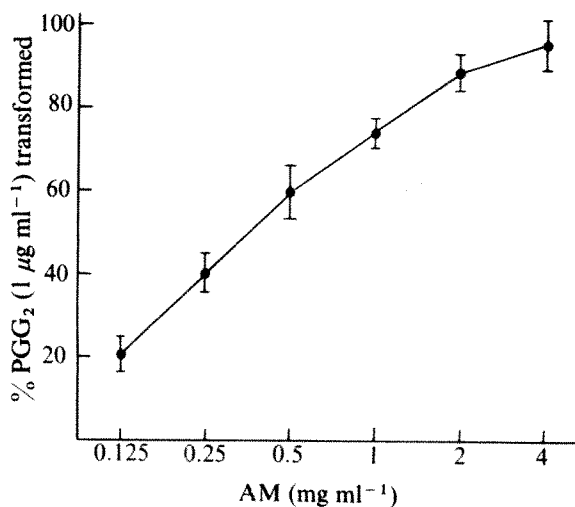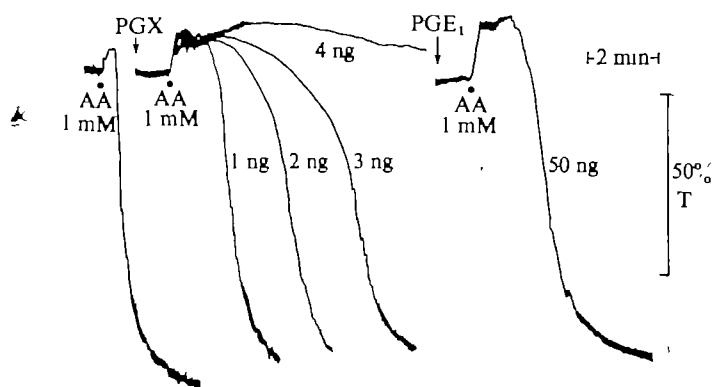
Fig. 4 The tracings show changes in light transmission through platelet-rich human plasma in a Born aggregometer. Comparison of anti-aggregatory potencies of PGX and PGE₁. PGX was obtained by incubation of 100 ng PGH₂ with 500 µg of aortic microsomes in 100 µl 0.05 M Tris buffer (pH 7.5) for 2 min at 22 °C and then stored on ice PGX and PGE₁ (10 µl) were added to human platelet-rich plasma 1 min before arachidonic acid (AA 1 mM). In this experiment PGX was at least 25 times more potent as an anti-aggregatory agent than PGE₁. Doses of PGX (1–4 ng) or PGE₁ (50 ng) are shown at the sides of the tracings.

PGF$_{2\alpha}$ and PGD$_2$ were not substrates for aortic microsomes, so that 15-keto PGs or other products of PG catabolism could not be considered as PGX. PGX is also unlikely to be a 15-hydroperoxy PG, because, firstly, 15-OOH PGE$_2$ contracts rabbit aorta[4], and, secondly, the product(s) of the spontaneous decay of PGX when bioassayed did not behave like PGE$_2$, PGF$_{2\alpha}$ or PGD$_2$. Lack of increased malondialdehyde formation from PGG$_2$ and PGH$_2$ excluded their enzymic transformation by aortic microsomes to HHT.

The presence of prostaglandin D$_2$ isomerase in homogenates of several tissues has been described[5]. Prostaglandin D$_2$ is a potent inhibitor of platelet aggregation[19]. PGX is, however, not PGD$_2$ because PGX is unstable and is a more potent anti-aggregatory agent than PGD$_2$. Besides, PGD$_2$ isomerase is present in the 100,000g supernatant[5], a fraction that in our experiments did not produce PGX from endoperoxides. Moreover, PGD$_2$ isomerase needs glutathione as a cofactor[5] and our incubations were carried out in the absence of cofactors.

The prostaglandin endoperoxides (PGG$_2$ and PGH$_2$) aggregate platelets[10,11] and this activity is probably mediated, wholly or partly, by the enzymic conversion of endoperoxides by platelet microsomes to the less stable TXA$_2$ (refs 6 and 7). Blood vessel microsomes transform endoperoxides to another unstable principle (PGX) which, contrary to TXA$_2$, has potent anti-aggregatory properties and relaxed, rather than contracted, vascular strips. Thus a balance between the amounts of TXA$_2$ formed by platelets and PGX formed by vessels might be critical for thrombus formation. Indeed, in the light of the discovery of this anti-thrombotic property associated with arterial walls, it is interesting to recall the

pre-Lister vitalistic view that in some way the arteries kept the blood fluid. Certainly, platelets adhere easily to any particle or surface, with the unique exception of the vascular endothelium. Generation of PGX by vessel walls could be the biochemical mechanism underlying their ability to resist platelet adhesion. Indeed, platelets attempting to stick to vessels may be releasing PGG$_2$ or PGH$_2$ which is then used by the vessel to generate PGX which prevents the platelets from sticking. Plaque formation on the arterial wall could hinder access of platelet endoperoxides to the PGX-generating system (or of any continuously formed PGX to the platelets) thereby enabling platelet deposition to occur.

PGX may also have other properties. It is well known that platelets repair lesions in the vascular endothelium[20]. An infusion of a small number of platelets is sufficient to restore vascular integrity in the damaged vessels of thrombocytopenic animals[21,22]. Studies of ultrastructure have revealed that platelets can be assimilated into the vascular endothelium or even incorporated into endothelial cells[23,24]. Biochemical cooperation between platelets and vascular endothelium in the generation of PGX could well contribute to the repair of vascular endothelium.

The fact that PGX is locally generated in the arterial wall and can relax arterial smooth muscle also suggests that it plays a part in the local control of vascular tone. A deficiency of PGX could, therefore, underly some forms of hypertension. It will be interesting to find out whether other tissues can also generate PGX and what other biological properties are associated with this class of compound.

We thank Dr F. Ubatuba for preparing the prostaglandin endoperoxides.

1 Hamberg, M., and Samuelsson, B, *Proc natn Acad Sci U S A*, 70, 899–903 (1973).
2 Nugteren, D H., and Hazelhof, E., *Biochim. biophys Acta*, 326, 448–461 (1973)
3 Hamberg, M, and Samuelsson, B, *Proc. natn. Acad Sci U S A*, 71, 3400–3404 (1974)
4 Van Dorp, D A, *XXIVth Int Cong Pure appl Chem*, 2, 117–136 (1974).
5 Hamberg, M, Svensson, J, and Samuelsson, B, *Proc natn. Acad Sci U S A*, 72, 2994–2998 (1975).
6 Needleman, P, Moncada, S, Bunting, S, Vane, J. R., Hamberg, M, and Samuelsson, B, *Nature*, 261, 550–560 (1976)
7 Moncada, S, Needleman, P, Bunting, S., and Vane, J. R, *Prostaglandins* (in the press).
8 Piper, P. J, and Vane, J R, *Nature*, 223, 29–35 (1969).
9 Palmer, M A, Piper, P J, and Vane, J. R, *Br. J Pharmac.*, 49, 226–242 (1973).
10 Svensson, J., Hamberg, M, and Samuelsson, B, *Acta Physiol Scand*, 94, 222–228 (1975)
11 Bunting, S, Moncada, S, and Vane, J R, *Br. J. Pharmac*, 57, 462P–463P (1976).
12 Malmsten, C, Hamberg, M, Svensson, J and Samuelsson, B, *Proc natn. Acad. Sci U S A*, 72, 1446–1450 (1975).
13 Lowry, O H, Rosebrough, W. J, Farr, A L, and Randall, R. J, *J biol Chem*, 193, 256–257 (1951)
14 Vane, J R, *Br J Pharmac*, 23, 369–373 (1964)
15 Gilmore, N, Vane, J. R, and Wyllie, J H, *Nature*, 218, 1135–1140 (1968)
16 Bunting, S, Moncada, S, Needleman, P, and Vane, J R., *Br J Pharmac*, 56, 334P–345P (1976)
17 Flower, R. J, Cheung, H S, and Cushman, D W, *Prostaglandins*, 4, 325–341 (1973)
18 Born, G V R, *Nature*, 194, 927–979 (1962)
19 Smith, J B, Silver, M J, Ingerman, C. M, and Kocsis, J J, *Thromb Res*, 5, 291–299 (1974).
20 Marcus, A J, *New Eng J. Med.*, 280, 1213–1220 (1969).
21 Roy, A J, and Djerassi, J, *Proc Soc exp. Biol Med.*, 139, 137–142 (1972)
22 Aursnes, I, *Acta physiol scand*, 88, 392–400 (1973)
23 Johnson, S. A, in *The Circulating Platelet* (edit by Johnson, S. A), 284 (Academic, New York, 1971)
24 Kitchens, C S, and Weiss, L, *Blood*, 46, 567–578 (1975)

letters to nature

A secular relativistic change in the period of a binary pulsar

THE discovery of the first binary radio pulsar PSR1913+16 by Hulse and Taylor[1] stimulated an interest in using relativistic effects to calculate more details of this system.[1–4] It was shown, for example, that observations of the apsidal motion could

allow the masses of the pulsar and its companion to be calculated. There are many mechanisms which could cause a secular change in the observed period of a binary pulsar (for example, mass exchange magnetic friction, electromagnetic radiation). Here we would like to discuss another possibility related to a change of orbital parameters of a pulsar in a binary system and we also estimate the effect of the emission of gravitational radiation on the period of the binary.

Let the intrinsic period of a pulsar be P_0, its orbital velocity v and the mass of the central star M_0; then the observed period P, neglecting the relative motion between the centre of mass of the binary system and the observer, is given by

$$P = \frac{P_0}{(1-(v/c)^2)^{1/2}(1-(2GM_0/c^2a))^{1/2}} \qquad (1)$$

where a is the semi-major axis of the relative orbit, G the gravitational constant and c the velocity of light. For simplicity we will assume that the pulsar moves in a circular orbit. Using Kepler's law to calculate the orbital velocity of the pulsar and keeping only first-order relative corrections, we get

$$P = P_0\left(1 + \frac{GM_0}{c^2a}\frac{3M_0 + 2M_p}{2(M_0 + M_p)}\right) \qquad (2)$$

where M_p denotes the mass of the pulsar.

Let T be the period of the binary system, then

$$\langle P \rangle = \frac{1}{T}\int_0^T P\,dt$$

is the mean period of the pulsar over the period of the binary system. Assuming that there is no mass exchange and M_0 and M_p are constant for the relative time variation of the period we obtain

$$\frac{\dot{P}}{\langle P \rangle} = -\frac{3M_0 + 2M_p}{2(M_0 + M_p)}\frac{GM_0}{c^2a}\frac{\dot{a}}{a}$$

From Kepler's law it follows that when the masses of the components are constant, the relative time variation of the semimajor axis is related to the time variation of the period of the binary system by

$$\frac{\dot{a}}{a} = \frac{2}{3}\frac{\dot{T}}{T} \qquad (4)$$

so finally we have the relationship

$$\frac{\dot{P}}{\langle P \rangle} = -\frac{3M_0 + 2M_p}{3(M_0 + M_p)}\frac{GM_0}{c^2a}\frac{\dot{T}}{T} \qquad (5)$$

Recently, Taylor *et al.*[5] obtained more accurate values of P and T for the pulsar PSR1913+16, giving $\dot{P}/\langle P \rangle = 1.49 \times 10^{-16}$ and $\dot{T}/T = 7.1 \times 10^{-15}$. Substituting probable values of masses $M_0 = 1.41M_\odot$ and $M_p = 1.41M_\odot$ and taking $a = 9.89 \times 10^{10}$ cm, for a ratio between $\dot{P}/P$ and $\dot{T}/T$ we obtain 1.75×10^{-6}. It therefore follows that for the pulsar PSR1913+16 the change in the period is caused by some mechanisms other than the change in orbital parameters.

Using equation (5), in principle, one should be able to detect the change in the period of a binary pulsar caused by the emission of gravitational waves. Landau and Lifshitz[6] calculated the change in semimajor axis of a binary system caused by the emission of gravitational radiation and obtained

$$\frac{\dot{a}}{a} = -\frac{64G^3 M_0 M_p(M_0 + M_p)}{5a^4c^5}$$

substituting this in (3) we have

$$\frac{\dot{P}}{\langle P \rangle} = \frac{32G^4 M_0^2 M_p(3M_0 + 2M_p)}{5c^7a^5} = \frac{2}{5}\left(\frac{R_g}{a}\right)^2\frac{r_g}{a}\left(3\frac{R_g}{a} + 2\frac{r_g}{a}\right)\frac{c}{a} \qquad (7)$$

where R_g is the gravitational radius of the central star and r_g is that of the pulsar. For PSR1913+16 we get $\dot{P}/\langle P \rangle = 1.9 \times 10^{-22}$ s^{-1}, a value which is unobservably small. In a favourable situation, however, with more massive stars forming a compact binary system, the change in the period of the pulsar should be observable. Let us take, for example, a binary system of two stars, one of which could be a black hole, of masses $M_0 = 15M_\odot$ and $M_p = 2M_\odot$ moving in a circular orbit with $a = 2 \times 10^{10}$ cm, then $\dot{P}/\langle P \rangle = 6.25 \times 10^{-16}$ s^{-1} which should be observable with present-day technology.

N.I.S. thanks the University of Warsaw for hospitality.

MAREK DEMIANSKI

Institute of Theoretical Physics,
University of Warsaw,
Warsaw, Poland

N. I. SHAKURA

Sternberg Astronomical Institute Moscow, USSR

1 Hulse, R. A., and Taylor, J. H., *Astrophys. J. Lett.*, 195, L51 (1975).
2 Brumberg, V. A., Zeldovich, Ya. B., Novikov, I. D., and Shakura, N. I., *Astr. Lett.*, 1, 5 (1975).
3 Blandford, R., and Teukolsky, S. A., *Astrophys. J.*, 205, 580 (1976).
4 Wagoner, R., *Astrophys. J. Lett.*, 196, L63 (1975).
5 Taylor, J. H., Hulse, R. A., Fowler, L. A., Gullahorn, G. E., and Rankin, I. M., *Astrophys. J. Lett.*, 206, L53 (1976).
6 Landau, L. D., and Lifshitz, E. M., *The Classical Theory of Fields* (Addison Wesley, Reading, Massachusetts, 1962).

Improved upper limits of gravitational deflection of polarised radiation

RECENTLY Harwit *et al.*[1] established upper limits for any polarisation dependence in the bending of radio waves by the Sun's gravitational field. Any differential angular splitting between orthogonal polarisations was shown by VLBI techniques, and from the constancy of the Stokes parameters in the carrier wave from a spacecraft, to be $< 10^{-3}$ and 10^{-6}, respectively, of the total bending. Since the latter technique is not sensitive to a splitting in the tangential direction, whereas interferometry is, development of both methods is important. These experiments form a unique test of the equivalence principle[1].

We report here further interferometer measurements. Fomalont and Sramek[2] have measured the total bending of microwave radiation by the Sun's gravitational field. In this paper we search for any polarisation splitting based on a separate analysis of Fomalont and Sramek's data taken with the four-element interferometer of the National Radio Astronomy Observatory (NRAO). For our analysis fringe phases in orthogonal circular polarisations were kindly provided by these investigators. They observed three small diameter radio sources, 0116+08, 0119+11, and 0111+02, at 2,695 and 8,085 MHz with the 35-km intermediate baseline interferometer, as the Sun passed close to their line of sight.

Although the baseline is nearly two orders of magnitude smaller than the baseline of our previous VLBI measurements, the present experiment has produced improved upper limits to the differential deflection. There are two main factors responsible for this improvement. First, the increased phase stability of the NRAO interferometer, together with the higher correlated flux densities at this angular resolution allow a significant increase in signal-to-noise ratio. Second, we were able to observe the sources at smaller solar elongations: 0116+08 was observed to within $< 1°$ of the Sun, 0119+11 to within nearly 3°, and 0111+02 to within $< 7°$.

The method of analysis was essentially the same as that employed in the VLBI experiment. A fringe phase difference, $\Delta\phi$, between the orthogonal circular polarisations, of the form

$$\Delta\phi = G(\rho)[(B_x \sin A - B_y \sin\delta \cos A)\sin H - (B_x \sin\delta \cos A + B_y \sin A)\cos H + B_z \cos\delta \cos A] + \text{constant}$$

was sought, where B_x, B_y, and B_z are the baseline components in wavelengths, A is the position angle of the splitting, H is the

Table 1 Upper limits to the polarisation splitting

		Tangential splitting		Radial splitting	
		Upper limits on K ($\rho = 1°$) (95% confidence)			
n	Frequency (MHz)	$K_{\text{antielockwise}}$ (10^{-3} arc s)	$K_{\text{clockwise}}$ (10^{-3} arc s)	K_{out} (10^{-3} arc s)	K_{in} (10^{-3} arc s)
1	2,695	70	95	70	60
	8,085	35	65	30	50
2	2,695	110	130	95	75
	8,085	90	70	60	60
3	2,695	165	195	90	50
	8,085	110	60	45	35

Upper limits to the polarisation splitting are given at a solar elongation of 1°. Clockwise tangential splitting is defined as that seen by an observer in which right circular polarisation appears rotated clockwise about the Sun away from left circular polarisation. In outward radial splitting right circular polarisation appears at greater solar elongation than left circular polarisation.

hour angle, and δ is the declination. Any angular splitting, $G(\rho)$, would be expected to depend on the solar elongation, ρ. Simple power law radial dependence of the form $G(\rho) = K\rho^{-n}$ ($n = 1, 2, 3$) was assumed.

No polarisation splitting was observed. Ninety-five per cent confidence upper limits were obtained by calculating χ^2 for various assumed values of the splitting. These upper limits are given in Table 1 for both frequencies, for both tangential and radial splitting, and the different values of n. Three intermediate length baselines were obtained by correlating signals from the 45-foot antenna with those from the three 85-foot antennae. These results were not statistically independent and the baseline yielding the smallest upper limits was chosen. The upper limits in Table 1 are ~50–100 marc s at a solar elongation of 1°.

With observations at low solar elongations, the present observations yield improved upper limits for any radial dependence as strong as ρ^{-3} or stronger. (The previous upper limits of Harwit *et al.*[1] were established for solar elongations of 5° and 7.8°, not 5 R and 7.8 R, as given in Table 1 of ref. 1.)

This work was supported by a grant from the Research Corporation. We extend special thanks to Drs E. Fomalont and R. Sramek of the NRAO for allowing us to carry out this analysis with their data. We thank Dr M. Harwit for valuable advice and discussions. The NRAO is operated by Associated Universities Inc., under contract with the NSF.

BRIAN DENNISON

Center for Radiophysics and Space Research,

JOHN DICKEY

National Astronomy and Ionosphere Center,
Cornell University,
Ithaca, New York 14853

DAVID JAUNCEY

CSIRO, Division of Radiophysics,
PO Box 76,
Epping, NSW 2121, Australia

[1] Harwit, M., Lovelace, R., Dennison, B., Jauncey, D., and Broderick, J., *Nature*, 249, 230 (1974).
[2] Fomalont, E., and Sramek, R., *Astrophys. J*, 199, 749 (1975).

Origin of Olympus Mons Escarpment by erosion of pre-volcano substrate

OF all the large shield volcanoes of Mars, Olympus Mons is unique in being fringed by a nearly continuous scarp rising 1–4 km above the surrounding terrain[1,2] (Fig. 1). We present evidence that the scarp was caused by preferential erosion of fragmental cratered terrain material previously adjacent to, and at present underlying the shield volcano. Removal of this material caused undercutting and collapse

of the shield flanks, producing the observed terraces and scarps.

A previous interpretation for the scarp was prompted by the recognition that basaltic lava flows which probably form the flanks of Olympus Mons are very resistant to any aeolian erosion[3]. King and Riehle[4] proposed that the volcano has a layered structure with basaltic lava flows capping a base of ash flow tuff deposits which would be readily erodable by aeolian processes. Such an origin for the scarp is unlikely, since volcanoes with basaltic tops and acidic bottoms are not produced typically by terrestrial igneous processes[5]. Rather, it is common for differentiation within magma chambers of terrestrial basaltic volcanoes to result in late stage eruptions of small volumes of more silicic lavas, as at Mauna Kea, Hawaii[6].

Undercutting such as apparently initiated the formation of the Olympus Mons scarp commonly occurs when an upper resistant layer is left unsupported through the removal of a less competent substrate. It is here proposed that the Olympus Mons substrate is a layer of fragmental rubble (megaregolith) resulting from repeated fracturing and brecciation of the surface by intense meteoritic bombardment. On the Moon, the cratered terrain megaregolith is 2–4 km in thickness[7], and a similar thickness may exist on Mars.

Much of the northern hemisphere of Mars is a smooth lava plain with a lower concentration of craters than the rest of the planet. We propose that this is not caused by any fundamental difference in crustal formation, but rather that the original cratered terrain has been stripped off in some erosional event, and the exposed surface veneered with lavas. It is not understood how this happened, nor why the destruction was limited to the northern hemisphere of Mars. Nonetheless, cratered terrain was removed leaving scattered remnants, and the edge of the cratered terrain is presently undergoing erosion[8]. The northern lowlands of Mars are 2–3 km lower than the southern cratered terrain. This difference in elevation is nearly the same as the height of the Olympus Mons scarp, and about the same as the lunar/martian megaregolith thickness. This implies that the erosional event removed the megaregolith only and did not excavate and remove large quantities of well-consolidated basement rock.

In this model the initial Olympus Mons lavas erupted upon cratered terrain, armouring it against later erosion. Removal of the cratered terrain megaregolith led to the undercutting of Olympus Mons lavas and to the formation of the scarp. Thus, the scarp itself is a relic of earlier erosion and not caused by the present erosional regime which only modifies the original structure. This would account for the previously puzzling contrast between the severe erosion of the scarp and the minor erosion higher up on the flanks of Olympus Mons. In some places (for example, the SW flank) later lavas from Olympus Mons flowed over the scarp, re-establishing the original slope. On
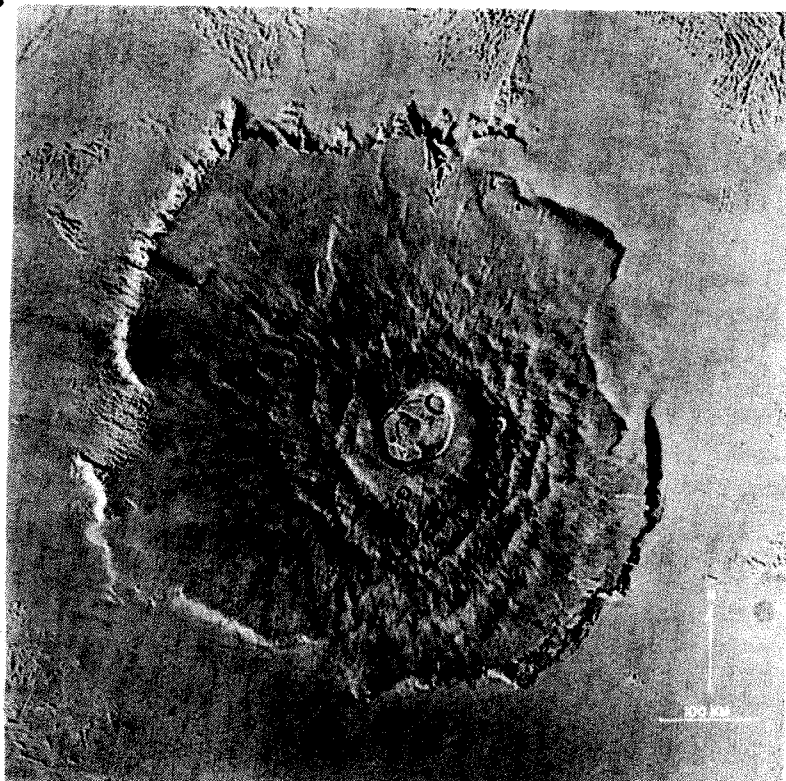
Fig. 1 Complex relationships between Olympus Mons shield volcano and the surrounding lava plain are indicated by the embayment of the eastern flank of the shield and by mantling of the south-west escarpment and plains by shield lavas. The photograph has been computer rectified, enhanced and mosaicked at the Image Processing Laboratory of the Jet Propulsion Laboratory by Karl Blasius and Jim Soha.

the other hand, it seems that lavas from the surrounding plain have lapped on to the eastern flank of the volcano. Thus, Olympus Mons spans the time interval from before the onset of the erosional event to a period more recent than the emplacement of the youngest lavas of the surrounding plain. From crater counts[9,10] this interval is estimated to have been $2 \times 10^8 - 2 \times 10^9$ yr, corresponding to a minimum average eruption rate of $\sim 10^{-3} - 10^{-2}$ km^3 yr^{-1}. For comparison, the eruption rate averaged over the past 70×10^6 yr of the Hawaii–Emperor chain is 1.5×10^{-2} km^3 yr^{-1} (ref. 11).

Neither the volcanoes of the Tharsis plateau nor those of Elysium have scarps such as bound Olympus Mons. According to the hypothesis developed here scarps would not form unless the ancient cratered terrain substrate was eroded away after the initial armouring eruptions. The geological map of the Tharsis region[12] shows 'hilly material', interpreted as remnants of cratered terrain, extending out from under the Tharsis volcanoes. This demonstrates that the cratered terrain was not completely destroyed during the erosion episode and thus, the Tharsis volcanoes were not undercut and did not develop scarps. Similar circumstances occur in the Elysium region. The Elysium dome is fringed by irregular rounded hills[12] which seem to be the erosional remnants of cratered terrain. Apparently, the ancient cratered terrain was not completely removed and thus scarps were not formed in this region.

JAMES W. HEAD
MARK SETTLE
CHARLES A. WOOD

Department of Geological Sciences,
Brown University,
Providence, Rhode Island 02912

1 Carr, M. H., *Scient. Am.*, **234**, 32–43 (1976).
2 Mutch, T. A., Arvidson, R. E., Head, J. W., Jones, K. L., and Saunders, R. S., *The Geology of Mars* (Princeton University Press, 1976).
3 Greeley, R., *Geology*, **1**, 175–180 (1973).
4 King, J. S., and Riehle, J. R., *Icarus*, **23**, 300–317 (1974).
5 Carmichael, I. S. E., Turner, F. J., and Verhoogen, J., *Igneous Petrology* (McGraw-Hill, New York, 1974.)
6 Macdonald, G. A., and Abbott, A. T., *Volcanoes in the Sea* (University of Hawaii Press, Honolulu, 1970).
7 Head, J. W., *Lunar Sci.*, **VII**, 354–356 (1976).
8 Sharp, R. P., *J. geophys. Res.*, **78**, 4073–4083 (1973).
9 Soderblom, L. A., Condit, C. D., West, R. A., Herman, B. M., and Kreidler, T. J., *Icarus*, **22**, 239–263 (1974).
10 Hartmann, W. K., *J. geophys. Res.*, **78**, 4096–4116 (1973).
11 Bargar, K. E. and Jackson, E. D., *J. Res. U.S. Geol. Surv.*, **2**, 545–550 (1974).
12 Carr, M. H., *Geologic Map of the Tharsis Quadrangle of Mars* (U.S. Geol. Surv. Misc. Invest. Map I-893, 1975).

Correlation of Martian surface heights with latitude of polar hood boundaries

A COMPARISON of Earth-based photography with topographic data from Mariner 9 reveals that the winter polar hoods on Mars tend to extend further towards the equator at lower topographic heights than at higher topographic heights. The correlation between latitude of the hood boundary and topographic heights is shown in Figs 1 and 2. In Fig. 1, contour lines derived from a map by Christensen[1] have been superimposed on the Lowell Observatory map of Martian Albedo Features and Topography[2]. To the author's knowledge, the Christensen map is the only published map that includes contour lines up to 65°N. The dashed lines represent the mean boundary of the polar hoods during twelve apparitions between 1905 and 1958, as determined from a study of Lowell Observatory's historical photographic collection. The southern hood boundary is in close agreement with a map by Capen[3] that was made from measurements of more recent apparitions (1962–68), while the northern boundary agrees well with Mariner imaging[4]. In Fig. 2, the latitude of the hood boundaries (at 10° intervals in longitude) is plotted as a function of topographic height. The dashed line was fitted to the southern latitude points by least squares. The north hood points were then shifted along the ordinate until their average value fell on the dashed line. An equally good, if not better, correlation is found if one plots the latitude of the boundary of the southern hood against atmospheric pressure from the work of Conrath *et al.*[5]. The hood is found to extend farther north over areas of higher pressure than over areas of lower pressure. Comparable atmospheric pressure data are not available for the Northern Hemisphere.

The explanation of the phenomenon reported here is not
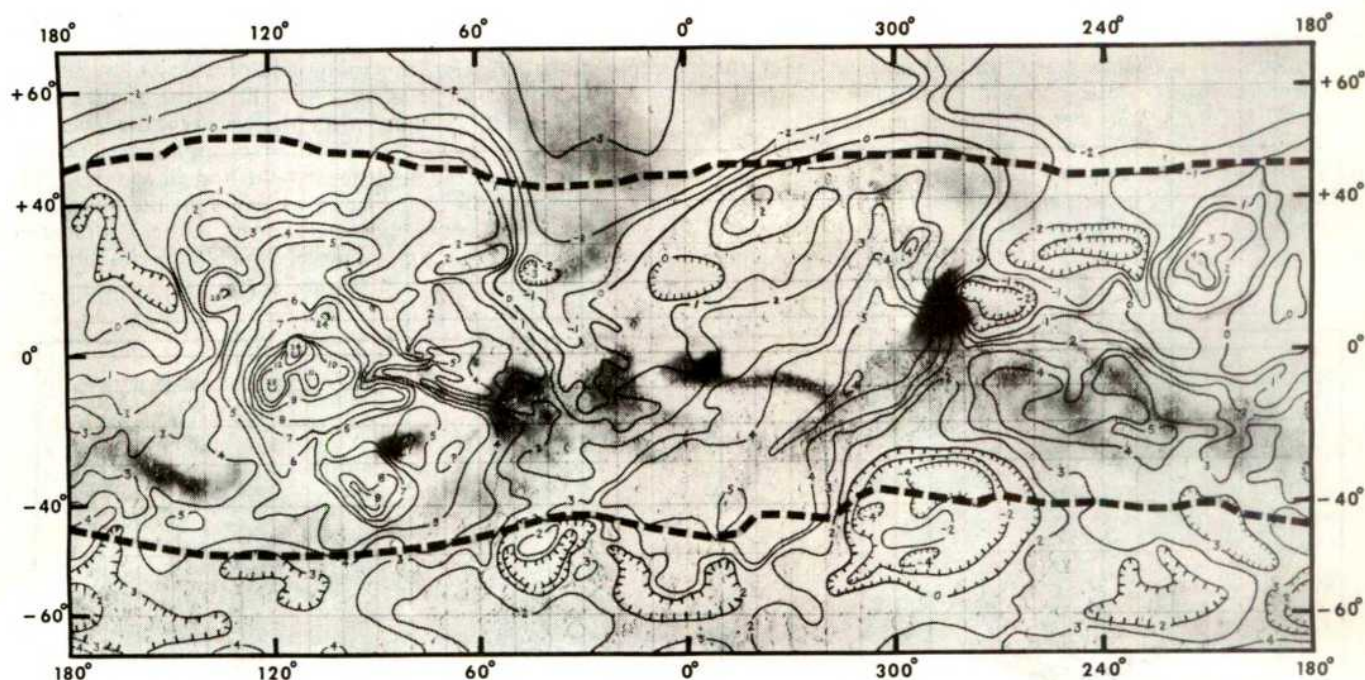
Fig. 1 Averaged boundaries for Martian polar hood clouds during northern and southern winters. The seasonal positions of Mars in its orbit that were used are L_s 315°–344° for the north and L_s 135°–164° for the south. The contour lines give heights in kilometres and have been redrawn by C. F. Capen from Christensen[1] on to the Lowell Observatory map of Martian Albedo Features and Topography[2].

Fig. 2 Correlation of surface heights with areographic latitude of the averaged boundaries of Martian polar hood clouds, measured on red- and green-filter photographs from the Lowell collection, taken from 1905 to 1958. The latitudinal scales have been offset to compensate for systematic differences between northern and southern winter hoods. The dashed line was fitted by least squares to the southern hood data. ×, North hood (L_s = 315°–344°); ○, South hood (L_s = 135°–164°).
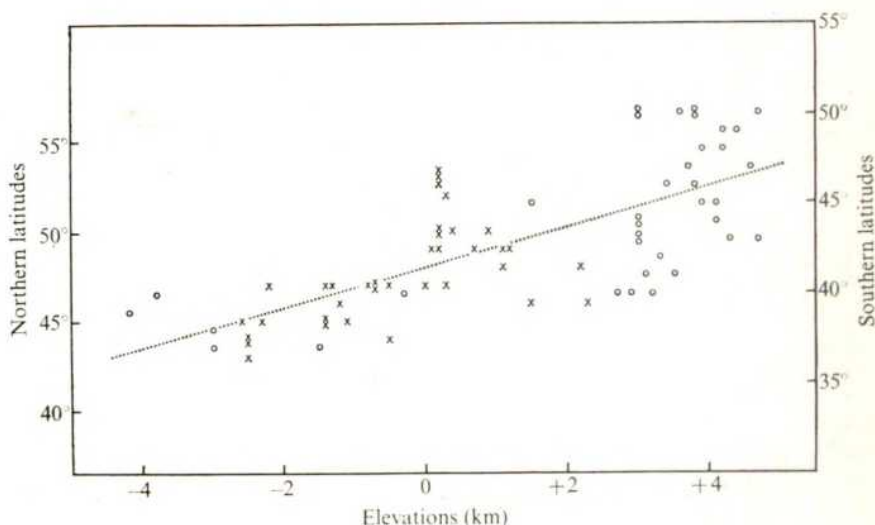
yet completely clear. Briggs and Leovy[4] have, however, argued that the hood clouds in the zone from 40 to 50°N must be water ice, while Farmer[6] has suggested that the bulk of the water vapour will be found in the lowest atmospheric layers.

LEONARD J. MARTIN

Planetary Research Center,
Lowell Observatory,
Flagstaff, Arizona 86001

1 Christensen, E. J., *J. geophys. Res.*, **80**, 2911 (1975).
2 Inge, J. L., *Mercury*, **2**, No 6, 10/11 (centrefold) (1973).
3 Capen, C. F., in *Mars Scientific Model*, Sec. 4.2, 26 (JPL Document 606–1, Pasadena, Calif., 1972).
4 Briggs, G. A., and Leovy, C. B., *Bull. Am. met. Soc.*, **55**, 278–296 (1974).
5 Conrath, B., *et al.*, *J. geophys. Res.*, **78**, 4267 (1973).
6 Farmer, C. B., *Icarus*, **28**, 279–289 (1976).

A granite cliff deep in the North Atlantic

WE report here preliminary results from the detailed sampling of granitic basement at a depth of 4,000 m at the lower edge of the Armorican continental margin near 48°N and 12°W. We consider that we have sampled here the extreme lateral limit of the continental crust where local exposures emerge from under a cover of Cainozoic, Mesozoic and probable Palaeozoic sediments.

The Goban Spur[1], part of the deep continental margin off France and the UK, lies between the Celtic shelf and the Porcupine abyssal plain ~ 550-km West of Brest and Land's End (Fig. 1). A seismic profile of the continental slope in this area shows a succession of three main 'gradins' (steps) with intermediate sedimentary basins (Fig. 2). The westernmost and deepest of the steps has a steep southwestern slope Granite Cliff 4,000 near the 4,000-m isobath.

• Twenty km further to the south-west is the marginal hill or high named by us Menez Bihan. At each of the three steps the substratum appears as an outcrop, passing under the sedimentary cover of the adjacent basin. The substratum seems to be continuous between the steps; it is 3-s d.t.t. (double travel time) deep to the east of the first step and 8-s d.t.t. deep at the foot of the continental slope (Fig. 2).

In December 1975, with the RV Le Suroït, we dredged fragments of granite from the steep slope we now call Granite Cliff 4,000 (Fig. 2). On the side of King Arthur

The dip of the slope of Granite Cliff 4,000 reaches at least 30° (the maximum slope which can be detected with the conventional 12-kHz precision depth recorder we used). The base of the slope is very straight in the studied area and this linearity, coupled with the apparent steepness of the slope, suggests that it is a fault escarpment. The foot of the slope is 4,150-m deep and the top lies at 3,150 m. Photographs of the slope show sediments on the lower and upper parts. In between, the sediment cover is thin and there are round outcrops, some of which show fractures.
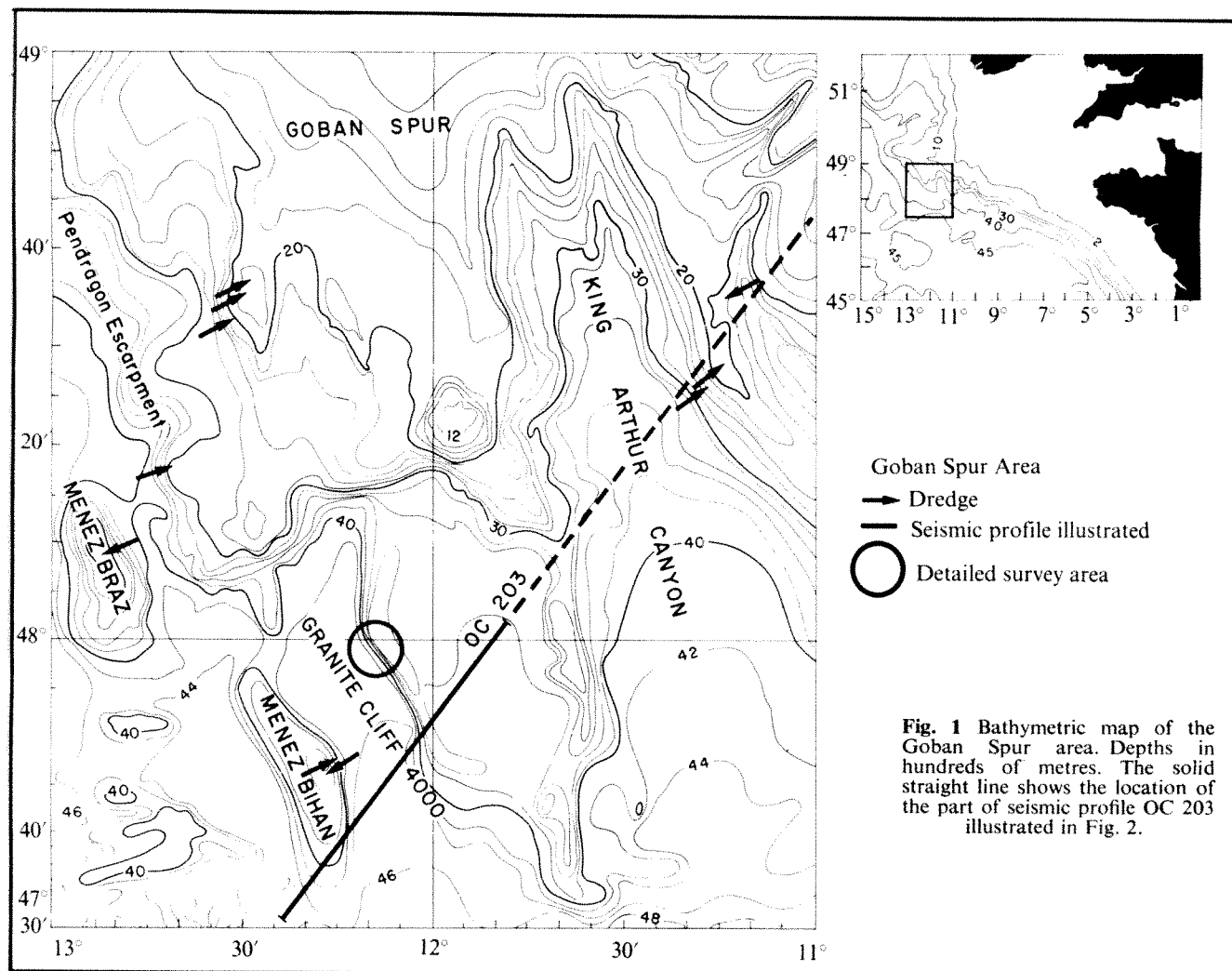
Fig. 1 Bathymetric map of the Goban Spur area. Depths in hundreds of metres. The solid straight line shows the location of the part of seismic profile OC 203 illustrated in Fig. 2.

Canyon, at a depth of 3,000 m, we dredged a pebble containing fossil material identified as Devonian. To investigate further these findings, and to make sure that the granite dredged was not erratic, we went back to the area in February 1976 with the RV Jean Charcot.

We made a bathymetric survey to clarify the relationship between Granite Cliff 4,000 and Menez Bihan and then moored three acoustic beacons on the bottom: two at the foot of the cliff and one near the top (all within the circle in Fig. 1). The distance between beacons was ~ 6,000 m. On board, on a plotting table, the track of the ship and the track of the dredge (or camera) on the bottom were then followed in real time with a precision of ~ 30 m. In this manner it is possible to visualise the work of the dredge on the bottom and, in combination with the tensiometer, to determine with precision the position of breaking-off of the rocks. For the seafloor photography, we used both a troika (a toboggan that slides on the bottom) and a camera frame attached to the cable just ahead of the dredge.

Four dredgings with the acoustic navigational control were made and in each dredge we found blocks of granitic rock. Their appearance and the dredge-wire tensiometer records indicate that some of the blocks were torn off outcrops, and support the evidence from the seafloor photographs that we are not dealing with ice-rafted erratics, except perhaps for a few associated pieces of dark or light granulite, the autochonous origin of which is still in doubt. In addition, the presence of all the rocks found can be reasonably associated with the setting in which they occur.

The granitic blocks are more or less weathered and are partially covered with ferromanganese crusts up to 12-cm thick. They are granodiorites, tending to quartz diorites. The minerals present are quartz with undulating extinction, biotite (abundant), andesine (abundant, automorphic and zoned), potassic feldspar (rather scarce), apatite, zircon and opaque minerals (J. Didier, personal communication). The fabric is clearly directional, and the biotites are slightly curved.

Menez Bihan

Granite Cliff 4,000
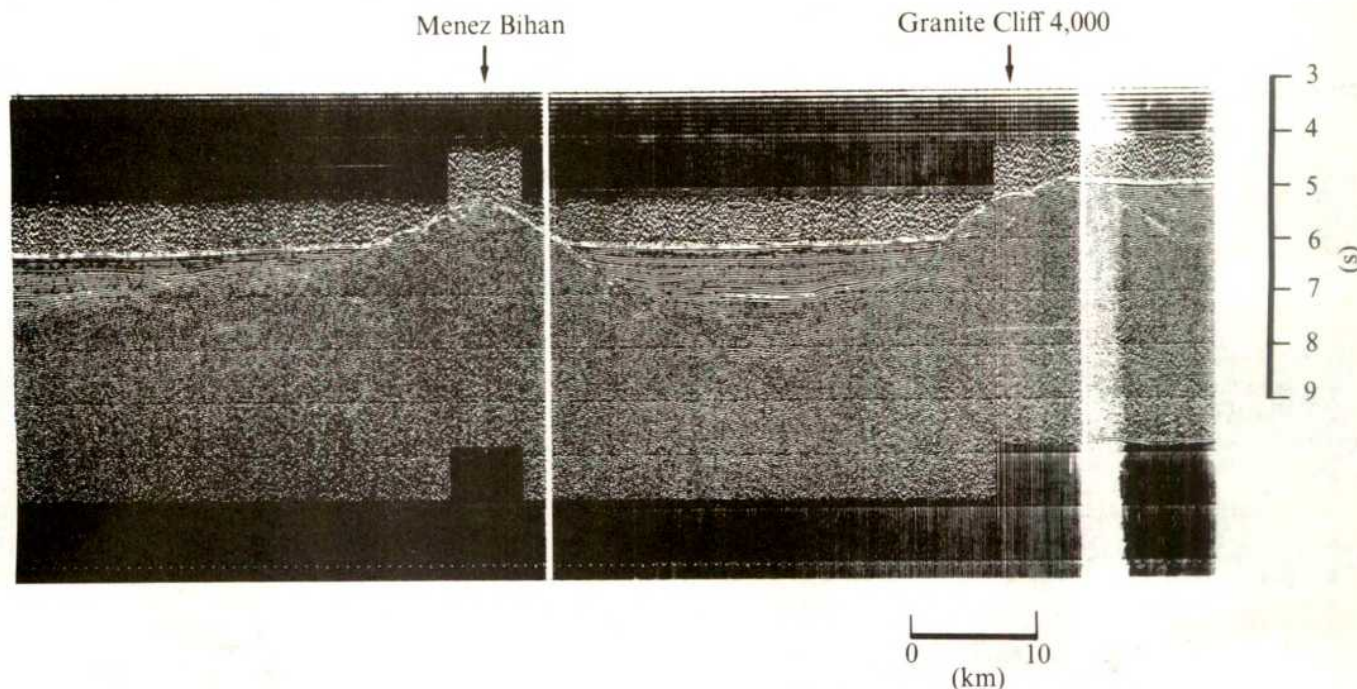
Fig. 2 Seismic profile OC 203 made by the RV Florence. Location of the profile is shown on Fig. 1. Vertical exaggeration is ×5.

Similar rocks possibly a little more altered, were dredged on the flanks of Menez Bihan (Fig. 2). The apparent continuity between Menez Bihan and Granite Cliff 4,000 (Fig. 2) is thus demonstrated. On geomorphic grounds (Fig. 1) it seems probable to us that Menez Braz to the north-west is also granitic, but we obtained no samples because of the considerable thickness of sediment there.

In three areas we dredged fragments of old sedimentary cover lying above the granitic substratum. From the flanks of Menez Bihan we collected a fragment of quartzitic sandstone. The remainder of the dredge-haul was entirely granitic Granite Cliff 4,000 yielded numerous samples of limestones some of them of possible Mesozoic age. On the Pendragon escarpment (Fig. 2), at a depth of between 3,000 and 2,000 m, the dredge recovered no granite, but a few metamorphic rocks and a rich sample of old sedimentary rocks: fine green sandstones, marly chalk with veins of calcite, grey claystone, sedimentary breccias and quartzite. No fauna has been found on preliminary examination, but we speculate that the rocks represent Palaeozoic or even older sediments.

The dredging of 'in place' granitic rocks at 4,000 m is exceptional and has important implications. In geometric reconstructions of the early opening of the North Atlantic, for example, many authors[2] have used a shallower depth (1,000 to 2,000 m) and hence a more easterly limit, for the edge of the continental crust. The maximum depth of the granite is evidently even greater than 4,000 m. Judging from the seismic profile (Fig. 2), the probable continuity of the granitic substratum between Granite Cliff 4,000 and Menez Bihan lies at a depth of 8 s d.t.t. (between 6,500 and 7,500 m).

To the south-west of Menez Bihan lies a low hill above the 4,000-m water depth (Fig. 1). Seismic profiles show that the basement of this hill is rough: quite different in character from that of Menez Bihan. Magnetic records show that the east flank of the low hill marks the beginning of a sharp anomaly (500×10^{-9} T) whereas the magnetic profile on the continental side of the hill is relatively flat. According to C. Williams (personal communication), anomaly 34 (~ 85 Myr) passes close by. The hill does not seem to have the acoustic

character of an intrusion, and may indeed mark the contact between oceanic crust and continental crust[3].

The granodioritic character of the granitic massifs we sampled suggests that they are orogenic granites, presumably of the intrusive type and of deep origin. It is difficult to envisage their emplacement during early rifting (90–100 Myr). None of the granites is alkaline and they are neither porphyritic nor strongly deformed. On the other hand, they may well be late generation (post-phase 2) granodiorites of Hercynian age (J. Didier, personal communication). They may represent intrusions in a granulitic basement which would explain the rocks belonging to the granulitic facies that we dredged from Granite Cliff 4,000 and from Menez Bihan. The basement underlying the continental slope itself is probably not made up of granodiorites. The latter usually appear as intrusions in the Hercynian chains (for example, the Armorican massif).

The oldest sedimentary rocks are epimetamorphic with at least two phases of deformation (folding). They resemble the Palaeozoic and Upper Precambrian (Brioverian) rocks of the Armorican massif (M. Gravelle, personal communication). We interpret this evidence as further support for arguing that the basement underlying the continental margin east of Menez Bihan is probably the older Hercyninan basement (granulite with granodiorite intrusions).

The direction of Granite Cliff 4,000 is similar to the Pendragon Escarpment (N 30° W) and may be related to Hercynian faults[4]. During the rifting phase of the North Atlantic, normal faulting associated with the subsidence of the continental blocks probably reactivated these old features. During the initial uplift phase, Menez Bihan would—in most models—have been on top of the uplifted crust and would consequently have been deeply eroded. Thus it seems reasonable that we dredged granite and Mesozoic limestone on Granite Cliff 4,000 and no Palaeozoic sediments. The upper part of Pendragon Escarpment, further away, was not, however, eroded to such an extent. Subsequently, thermal processes led to inversion of relief, with normal faulting along the Hercynian zone of weakness and subsidence of the continental margin.

• David Needham and Xavier Le Pichon gave us useful advice in their reviews of the manuscript. Jean Didier, Michel Gravelle, Claude Babin and Annick Pelhate made preliminary descriptions of the samples.

GUY PAUTOT
VINCENT RENARD
GERARD AUFFRET
LEO PASTOURET

Centre Océanologique de Bretagne,
B.P. 337, 29273, Brest Cedex,
France

OLIVIER DE CHARPAL

Institut Français du Pétrole,
1 à 4, avenue de Bois-Préau,
92500 Rueil-Malmaison,
France

1 Laughton, A. S., Roberts, D. G., and Graves, R., *Sheet 3, No. C 6568* (Institute of Oceanographic Science, United Kingdom, 1975).
2 Bullard, E. C., Everett, J. E., and Smith, A. G., *Phil. Trans. R. Soc.*, **A1088**, 41–51 (1965).
3 Pitman, W. C., III, Larson, R. L., and Herron, E. M., *The Age of the Ocean Basins.* (The Geological Society of America, 1974).
4 Roberts, D. G., *Geology of Continental Margins* (edit. by Burk, C. A., and Drake, C. L.), (Springer-Verlag, New York, 1974).

Infrared polarisability of hexagonal ice

THE limiting high frequency polarisability, α, of a solid is the algebraic sum of the optical, or electronic, polarisability, α_e, of the ions or molecules, and the infrared polarisability, α_{IR}, caused by the displacement, or vibration, of ions or molecules within the lattice. For most solids both α_e and α_{IR} are independent of volume and temperature, but for a number of ionic solids and for polymorphs of ice, α_{IR} is not so. We describe the temperature dependence of α_{IR} of hexagonal ice, and discuss it in terms of contributions from anharmonic effects.

For an isotropically polarisable material, α is given by the Clausius–Mossotti relationship

$$\frac{4}{3}\pi N\alpha = V\left(\frac{\varepsilon_\infty - 1}{\varepsilon_\infty + 2}\right) \qquad (1)$$

where N is Avogadro's number, V the molar volume and ε_∞ the limiting high frequency relative permittivity of the orientation polarisation. V for H_2O ice was calculated for a density of 0.9169 Mg m^{-3} at 270 K, and the average expansivity[1-5] is given in Fig. 1. V for D_2O ice was calculated for a density of 1.0175 Mg m^{-3} at 270 K and Dantl's[3] expansivity data. ε_∞ of H_2O ice for 2–77 K is taken from Gough's data[6] and for 77–273 K from our own[7]. ε_∞ of D_2O ice is known only above 77 K (ref. 8).

α_e for an H_2O molecule in ice is 1.5 Å^3 (ref. 9), and α_e for a D_2O molecule in the liquid phase, calculated from the Lorentz–Lorenz equation, is 1.46 Å^3 (ref. 8), which value is likely to be the same in D_2O ice. α_{IR} ($= \alpha - \alpha_e$) calculated from these values of α and α_e for both the ices is plotted against temperature in Fig. 2, which shows that the infrared polarisability of ice increases with temperature.

$$\alpha_{IR} = \alpha_{0,IR} + AT^2 \qquad (2)$$

where $\alpha_{0,IR}$ is the infrared polarisability at O K and A is an empirical constant. The values of $\alpha_{0,IR}$ and A are 1.64 Å^3 and 2.03 × 10^{-6} Å^3 K^{-2} for H_2O and 1.49 Å^3 and 2.95 × 10^{-6} Å^3 K^{-2} for D_2O ice. The T^2 increase in α_{IR}, which seems unique to hexagonal ice, is attributable to the anharmonicity of mechanical and electrical vibrations of H_2O molecules, as a consequence of which, the force constant of molecular vibrations becomes amplitude and temperature dependent. $\alpha_{0,IR}$ can then
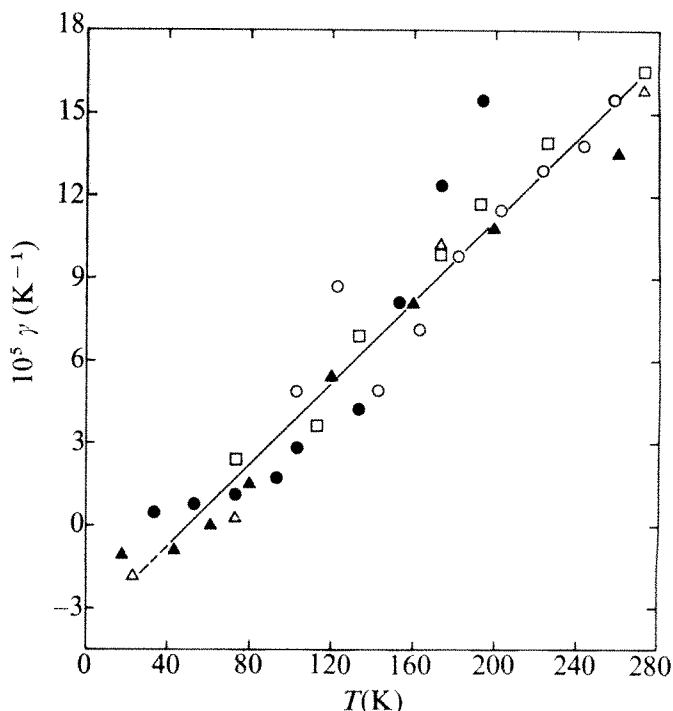
Fig. 1 The isobaric expansivity, γ, of H_2O ice plotted against temperature. The values were interplotted from, $\triangle$, Jacob and Erk[1]; ■, Powell[2], ▲, Dantl[3] for bulk expansion; and are as given by ○, LaPlaca and Post[4]; and ●, Brill and Tippe[5] for lattice constant expansion. The straight line between 30 and 273 K is the 'best' curve, in lieu of a least square fit which seems inappropriate because numerical data from various authors are lacking. Results from refs 2, 4 and 5 have been converted to values for polycrystalline ice.

be interpreted as the infrared polarisability if the lattice vibrations were entirely harmonic, as is anticipated at 0 K.

The anharmonicity of lattice vibrations is also reflected in the temperature dependence of heat capacity and expansivity, and a consideration of the latter is useful in an analysis of the T^2 dependence of α_{IR}. The temperature coefficient of α_{IR} at a constant pressure may be divided into two

$$\left[\frac{\partial \alpha_{IR}}{\partial T}\right]_P = \left[\frac{\partial \alpha_{IR}}{\partial T}\right]_V + \gamma\left[\frac{\partial \alpha_{IR}}{\partial \ln V}\right]_T \qquad (3)$$

where γ is the volume expansivity. $(\partial \alpha_{IR}/\partial T)_V$ constitutes the volume-independent temperature effect and $(\partial \alpha_{IR}/\partial \ln V)_T$ the temperature-independent volume effect.

Because of the lack of an adequate theory of anharmonic effects, the exact relation between γ and T is not known. For ice, such a relation is further uncertain because of the great variation between the reported values of its γ, as evident in Fig. 1. In the range, 30–273 K, γ can still be adequately described by the equation, $\gamma = \gamma_0 + mT$, as shown in Fig. 1, although the linear relationship is not anticipated to hold near 0 K because theory requires that γ should be zero at 0 K. For H_2O ice, $\gamma_0 = -3.6 \times 10^{-5}$ K^{-1} and $m = 7.39 \times 10^{-7}$ K^{-2}. Substituting for γ in equation (3), one obtains

$$\left[\frac{\partial \alpha_{IR}}{\partial T}\right]_P = \left[\frac{\partial \alpha_{IR}}{\partial T}\right]_V + \gamma_0\left[\frac{\partial \alpha_{IR}}{\partial \ln V}\right]_T + m\left[\frac{\partial \alpha_{IR}}{\partial \ln V}\right]_T T \qquad (4)$$

By differentiating equation (2)

$$\left(\frac{\partial \alpha_{IR}}{\partial T}\right)_P = 2AT \qquad (5)$$

A comparison of equations (4) and (5) suggests that, unless $(\partial \alpha_{IR}/\partial T)_V$ is a linear function of temperature

$$\left(\frac{\partial \alpha_{IR}}{\partial T}\right)_V = -\gamma_0 \left(\frac{\partial \alpha_{IR}}{\partial \ln V}\right)_T$$

and

$$\left(\frac{\partial \alpha_{IR}}{\partial \ln V}\right)_T = \frac{2A}{m}$$

For H_2O ice, using values of A and m given above, $(\partial \alpha_{IR}/\partial \ln V)_T = 5.49$ Å^3 and $(\partial \alpha_{IR}/\partial T)_V = 1.98 \times 10^{-4}$ Å^3 K^{-1}.

The contribution to $(\partial \alpha_{IR}/\partial T)_P$ in equation (3) from a temperature change alone, $(\partial \alpha_{IR}/\partial T)_V$, is positive for ice, as it is for ionic crystals of NaCl structure[10], but the contribution from a volume change, $\gamma(\partial \alpha_{IR}/\partial \ln V)_T$, is negative at $T < 50$ K and positive at $T > 50$ K, because γ changes sign near 50 K (Fig. 1). Near 273 K, $\gamma(\partial \alpha_{IR}/\partial \ln V)_T$ increases to nearly 4 $(\partial \alpha_{IR}/\partial T)_V$. Thus the difference between the respective temperature variation of α_{IR} at constant volume and constant pressure is substantial near 273 K and all of the increase in α_{IR} (and ε_∞) with temperature results from the associated volume change rather than to any explicit dependence on temperature.

As the nonlinear effects of anharmonic terms in the intermolecular potential, responsible for expansivity, are anticipated to disappear near 0 K, $\gamma \to 0$ as $T \to 0$ K. Therefore $(\partial \alpha_{IR}/\partial T)_V \to 0$ at 0 K as does $(\partial \alpha_{IR}/\partial T)_P$. A complete analysis of the interesting features of α_{IR} in the low temperature range would require accurate values of γ, $(\partial \alpha_{IR}/\partial T)_V$ and $(\partial \alpha_{IR}/\partial \ln V)_T$, not available at present. The validity of the relation between α_{IR} and T^2 for ice, however, is not likely to be significantly affected by a possible negative value of γ below 50 K, for the data points below 50 K lie in a relatively small range of T^2, as shown in Fig. 2.

The pressure dependence of α_{IR} and ε_∞ of ice can be predicted from

$$\left[\frac{\partial \alpha_{IR}}{\partial P}\right]_T = -\beta \left(\frac{\partial \alpha_{IR}}{\partial \ln V}\right)_T \qquad (6)$$

where β is the compressibility. For H_2O ice $\beta = 12 \pm 1$ Mbar^{-1} (ref. 11) at $20 < T < 273$ K, and therefore $(\partial \alpha_{IR}/\partial P)_T = -65.9$ Å^3 Mbar^{-1}.
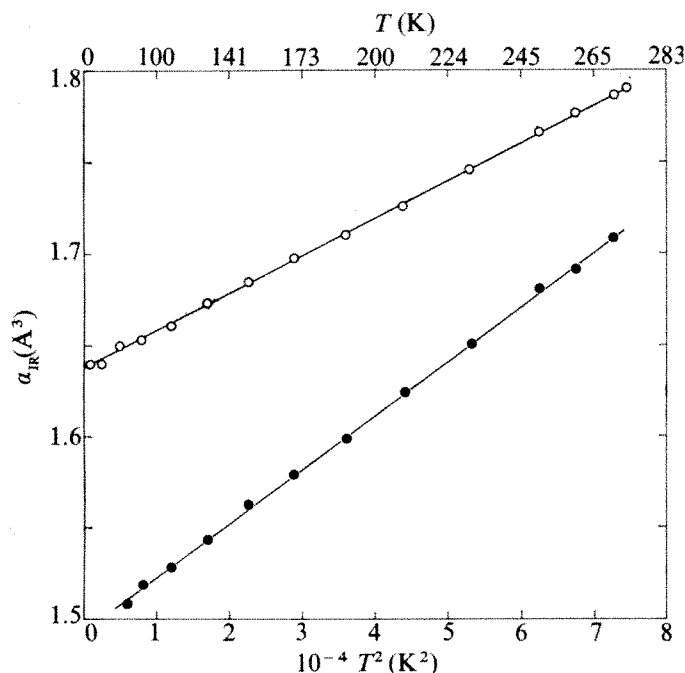
By differentiating equation (1) with respect to pressure, one obtains

$$\left(\frac{\partial \varepsilon_\infty}{\partial P}\right)_T = \frac{1}{3}(\varepsilon_\infty + 2)(\varepsilon_\infty - 1)\left[\beta + \frac{1}{\alpha_{IR}}\left(\frac{\partial \alpha_{IR}}{\partial P}\right)_T\right] \qquad (7)$$

At 270 K, $\varepsilon_\infty = 3.19$ (ref. 7), and therefore from equations (6) and (7), $(\partial \varepsilon_\infty/\partial P)_T = -94$ Mbar^{-1}. Thus, both the ε_∞ and α_{IR} of ice are anticipated to decrease by $\sim 3\%$ kbar^{-1} near 270 K.

γ (ref. 3) and β (ref. 11) for H_2O and D_2O ice are the same and at 270 K, $\varepsilon_\infty = 3.09$ (ref. 8) for D_2O ice. The analysis of α_{IR} for D_2O ice thus gives: $(\partial \alpha_{IR}/\partial \ln V)_T = 7.98$ Å^3, $(\partial \alpha_{IR}/\partial T)_V = 2.87 \times 10^{-4}$ Å^3 K^{-1}; $(\partial \alpha_{IR}/\partial P)_T = -95.8$ Å^3 Mbar^{-1} and $(\partial \varepsilon_\infty/\partial P)_T = -156$ Mbar^{-1}. The decrease in α_{IR} and ε_∞ on application of pressure is, therefore, anticipated to be $\sim 60\%$ more in D_2O than in H_2O ice.

The magnitude of the contribution to ε_∞ from vibrational polarisation is directly related to the negative second moment of the absorptivity, according to the Kramers–Kronig relationship (see ref. 7 for details). It follows that the application of pressure should either increase the frequency of the major absorption bands and/or decrease the absorptivity of the infrared radiation in ice.

The above conclusions are of significance because measurements of both ε_∞ and the infrared spectra of hexagonal ice with pressure are difficult, and because it is widely assumed that the application of pressure should increase ε_∞ of ice.

G. P. Johari
S. J. Jones

Glaciology Division,
Department of the Environment,
Ottawa, K1A OE7, Canada

1 Jacob, M., and Erk, S., *Z. ges. Kälteind.*, 35, 123–130 (1928).
2 Powell, R. W., *Proc. R. Soc.*, A247, 464–466 (1958).
3 Dantl, G., *Z. Phys.*, 166, 115–118 (1962).
4 LaPlaca, S. and Post, B., *Acta Crystallogr.*, 13, 503–505 (1960).
5 Brill, R., and Tippe, A., *Acta Crystallogr.*, 23, 343–345 (1967).
6 Gough, S. R., *Can. J. Chem.*, 50, 3046–3051 (1972).
7 Johari, G. P., *J. chem. Phys.*, 64, 3998–4005 (1976).
8 Johari, G. P., and Jones, S. J., *Proc. R. Soc.*, A349, 467–495 (1976).
9 Dorsey, N. E., *Properties of Ordinary Water Substances*, 177 (Hafner Publishing Co., New York, 1968).
10 Lowndes, R. P., and Martin, D. H., *Proc. R. Soc.*, A316, 351–375 (1970).
11 Leadbetter, A. J., *Proc. R. Soc.*, A287, 403–425 (1965).

T (K)

Fig. 2 The infrared polarisability, α_{IR}, of $\bigcirc$, H_2O and $\bullet$, D_2O ice plotted against the square of temperature. The relation between α_{IR} and T^2 is little affected by the variation in the reported values of γ seen in Fig. 1.

The valence of transition metal atoms in metallic alloys

THE 'valence' of transition metal atoms in metallic alloys has long been and still is a source of controversy. On the one hand, there are measurements such as those of Haworth and Hume-Rothery[1] which indicate quite small valences, there are electron concentration rules in numerous Hume-Rothery β and γ phases which indicate 'zero' valences, and there are measurements (for example, Coppens[2]) which even indicate 'negative' valences, whereas on the other hand, melting points, hardness and heats of atomisation indicate much higher valences. In this paper we point out that the valence of a transition metal atom in a metallic alloy cannot usefully be described by a single number. Three parameters are required: the number of outer electrons contributed to partly filled energy bands, the number entering filled bands lying below the Fermi level, and a parameter $(0 < \alpha < 1)$ describing the effectiveness for cohesion of the hybridisation of the filled bands originally derived from d states with those originally derived from s and/or p states.

This is because the valence electrons contributed by the transition metal may become part of two distinct groups which have different properties. The number and properties of electrons in one group—partly filled sheets of Fermi

surface (partly filled energy bands)—can be explicitly measured in suitable cases, whereas the properties of the other group—electrons in filled energy bands lying below the Fermi level—are always seen with those of the other group. Thus the valence of a transition metal atom in a metallic alloy cannot usefully be described by a single number. One parameter is required to specify the number of electrons contributed to partly filled energy bands in the alloy, another to describe those contributed to filled energy bands which have mainly d character. A further parameter, $0 < \alpha < 1$, is required to describe the effectiveness of the hybridisation of the filled bands originally derived from d states with those originally derived from s (and/or p) states, which ensures that the electrons are cohesive ($\alpha > 0$) even though they lie below the Fermi level. Such a description does not seem to be basically different to a valence bond description that considers the formation of hybridised bond orbitals which are derived from d, s and/or p states (see, for example, Pauling[3]; Altman, Coulson and Hume-Rothery[4]), but it does have the advantage when applied to metallic alloys of allowing one to recognise the difference between the numbers of itinerant electrons in partly filled energy bands and those in filled energy bands, which give the differences in physical measurements or estimates of valence referred to above.

Thus it is obvious that in determining the valence of Mn, Fe, Co and Ni from phase boundaries in certain alloys, Haworth and Hume-Rothery[1] were measuring properties controlled by the number of electrons contributed by the transition metals to the partly filled bands. In β and γ brass-type phases where many transition metals are said to be 'zero valent' to satisfy electron concentration rules, the outer electrons of the transition metal atoms are in filled bands below the Fermi level and the valence electrons contributed by the other component are sufficient to provide the required electron concentration in the partly filled bands. This is confirmed by the similarity of the Fermi surface of β-PdIn to those of β-AgZn and β-CuZn, indicating in each case, three electrons in partly filled sheets of Fermi surface in Brillouin zones one and two (Jan, Pearson and Saito[5]). Since indium contributes three valence electrons, the 10 outer electrons of Pd can be regarded as occupying filled bands below the Fermi level, and 'zero valence' is to be interpreted as contributing no net electrons to the conduction bands. Similarly an apparent negative valence of a transition metal is interpreted as retention of some of the valence electrons contributed by the other alloying component(s) in filled bands below the Fermi level, so that the electron concentration in partly filled sheets of Fermi surface (conduction bands) is reduced below the number of electrons contributed by the other alloying component(s).

The melting points of the transition metals and their heats of atomisation generally rise to a maximum in the fifth or sixth Group where the most effective cohesion from the outer electrons, is obtained. Although Fermi surface studies show the noble metals, Cu, Ag and Au, to have only one electron in partly filled bands, the high melting points (1,083 °C for Cu) and the heats of atomisation (81.1 kcalorie per gram atom at 298.15 K for Cu (Brewer[6])) indicate far more cohesion than that obtained from one outer electron (20–30 kcalorie per gram atom in heat of atomisation). This extra cohesion comes from hybridisation of the bands derived from d states with those derived from s and p states so that they contribute to cohesion even though they are full and lie below the Fermi level. The low melting point and the heat of atomisation of Zn (31.2 kcalorie per gram atom at 298.15 K (Brewer[6])), however, correspond to the two outer electrons in partly filled bands and indicate virtually no remaining hybridisation of the d bands to contribute to cohesion (that is $\alpha \sim 0$).

In summary, then, if V_i represents the number of outer electrons contributed by the transition metal to partly filled energy bands and V_1 the number contributed to filled energy bands, V_i may, in favourable cases, be determined from measurements of properties sensitive to the electron concentration in partly filled energy bands (Fermi surface, phase boundaries, etc.) and the number of valence electrons contributed by a non-transition element in the alloy. In considering properties such as melting points and cohesive energy, the contributions of both αV_1 and V_i are to be considered. Thus, for example, although electron concentration rules indicate $V_i \sim 0$ in β-brasses such as AlCo or AlNi, their high melting points (1,645 and 1,638 °C respectively) indicate a sizeable value of αV_1, which with the three electrons per formula unit contributed by Al, gives the large cohesive energy. The classical chemical valences of the transition metals (2+, 3+, etc.) are normally found in compounds with filled energy bands that are insulators or semiconductors.

W. B. PEARSON

Department of Physics,
University of Waterloo,
Waterloo, Ontario, Canada

1 Haworth, J. B., and Hume-Rothery, W., *Phil. Mag.*, 43, 613–629 (1952).
2 Coppens, P., *Xth International Congress, International Union of Crystallography, Amsterdam* (1975).
3 Pauling, L., *Proc. R. Soc.*, A196, 343–362 (1949).
4 Altman, S. L., Coulson, C. A., and Hume-Rothery, W., *Proc. R. Soc.*, A240, 145–159 (1957).
5 Jan, J.-P., Pearson, W. B., and Saito, Y., *Proc. R. Soc.*, A297, 275–287 (1967).
6 Brewer, L., in *High Strength Materials* (edit. by Zackay, V. F.), 12–103 (Wiley, New York, 1965).

The glass transition temperatures of phosphoric acids

THE vitrification of phosphoric acids is relatively easy, a glass often being formed simply on cooling from the liquid melt. Special procedures, however, are required to crystallise these materials. The glass transition temperature, T_g, is the temperature above which there is relatively rapid molecular motion; below T_g translational motion of molecules is inhibited. In standard texts[1] it is stated that the glass transition temperature of orthophosphoric acid, H_3PO_4, is -121 °C. The source of the determination is not always cited but it seems to be that of Kobeko *et al.*[2] who determined T_g from measurements of the temperature dependence of the electrical conductivity of phosphoric acid. Although published almost 40 yr ago there does not seem to have been a more recent determination. We present here evidence from broad-line nuclear magnetic resonance (NMR) studies that the estimate of -121 °C for the glass transition temperature of ortho-phosphoric acid is wrong. A discussion of the interpretation of the electrical conductivity of phosphoric acid would be inappropriate beyond observing that several different ionic or molecular processes may be responsible for the transfer of electrical charges. Thus, an unambiguous determination of a glass transition from such data may be far from straightforward.

Broadline NMR spectra are obtained when solids or viscous liquids are studied, and the linewidth, δH, can be interpreted as a measure of the extent of molecular motion. With the onset of molecular motion the linewidth is reduced; this is because of the averaging effect of local molecular motion on the magnetic field 'seen' by the nuclei, such as the protons in the phosphoric acids.

In Fig. 1 typical proton NMR, ^1H, spectra are given for a sample prepared by dehydration at 300 °C of analytical reagent grade phosphoric acid, 85–90% H_3PO_4. Experimental methods used for the measurement of the NMR spectra are similar to those used for the determination of the T_g of α-D glucose penta-acetate and epoxy resins[3,4].

In experiments such as these the NMR linewidth is determined by the spin–spin relaxation time, T_2, and the linewidth, δH, is inversely related to this relaxation time

$$\delta H = \frac{\text{const}}{\gamma T_2}$$

where δH, the linewidth, is measured between the maxima and minima of the first derivative of the absorption signal (Fig. 1) γ is the gyromagnetic ratio and the value of the constant depends on the line shape[3,4].
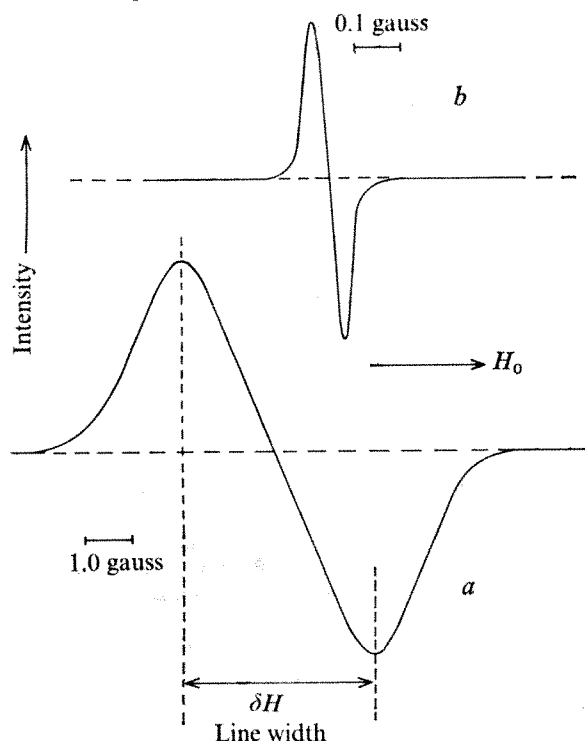
Fig. 1 First derivatives of NMR absorption spectra for phosphoric acid with $R = H_2O/P_2O_5 = 1.17$. *a*, Rigid glass at 169 K ($-104\,°C$). *b*, Fluid liquid at 302 K (29 °C). The linewidth was determined from at least six repeat spectra. The spectrometer modulation was $\sim 1/5\,\delta H$ to avoid modulation broadening of the spectra. Measurements taken from ref. 6.

The change in linewidth with temperature for a phosphoric acid glass forming system is given in Fig. 2. At low temperatures the linewidth is large because of dipolar broadening because of the presence of other nuclei with magnetic moments. Thus, the linewidth is essentially determined by the local magnetic fields, H_{loc}, which depend on the local arrangements of neighbouring protons. When the temperature is raised sufficiently to permit molecular motion, the time average local magnetic field, $\langle H_{loc}\rangle$, is reduced. From NMR theory it is shown that relaxation processes have longer T_2 relaxation times when molecular motion occurs, and hence the linewidth δH decreases when T_2 increases.

Several methods have been proposed for the determination of glass transition temperatures from the change of linewidth with temperature[8], all of which yield essentially the same estimate of T_g when the transition is 'sharp'. The inflection point of the δH against T curve (Fig. 2) was used to estimate T_g for these phosphoric acid glasses. The transition illustrated in Fig. 2 is very sharp.

For specification of glass transition temperatures it is essential to include the experimental procedure in the operational definition of T_g, as was pointed out by Gee $\sim 10\,\mathrm{yr}$ ago[5]. Maklakov and Pimenov[8] have discussed the relationships between NMR line narrowing and other methods of estimating T_g and McCall[9] has commented on NMR measurements of relaxation behaviour and the theories of glass transition behaviour. For the present purposes the inflection point
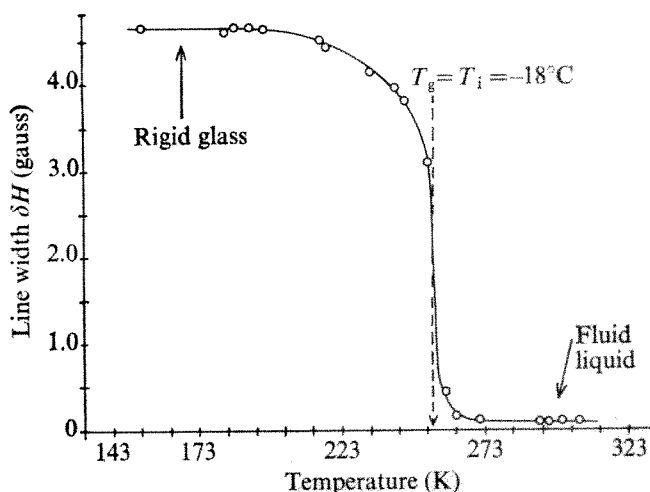
Fig. 2 NMR linewidth against temperature for phosphoric acid with $R = 1.33$. The glass transition temperature T_g is determined from the inflection point in the transition region[6].

of the NMR linewidth against temperature curve will be defined as the glass transition temperature, that is, $T_i = T_g$.

There are several theoretical empirical mixture rules for the correlation of the glass transition temperatures of two component systems[10], but none are established unequivocally. Thus, to represent the effect of composition of phosphoric acids on their glass transition temperatures it is convenient to plot T_g against R, where R is the molar ratio of H_2O/P_2O_5. Use of this compositional variable allows inclusion of metaphosphoric acid HPO_3, $R = 1$, and normal laboratory reagent 85–90% orthophosphoric acid, $R \simeq 4$. From Fig. 3, it is clear that there

Fig. 3 Glass transition temperatures against R. T_g was determined from NMR linewidth against temperature graphs as in Fig. 2. The straight line was fitted by a linear regression procedure with the assumption that R was error free and all the error is in T_g, its equation is $T_g = 281.1 - 17.28\,R$, $1 \leqslant R \leqslant 4$. The correlation coefficient is 0.99_3. Data from ref. 6, $R = 1.17$ and 1.33; ref. 7, $R = 2.10, 2.78$ and 3.95; $R = 2.18$ (unpublished data of B.E. and N. A. Miller), ▲., ref. 11; ●, ref. 2.
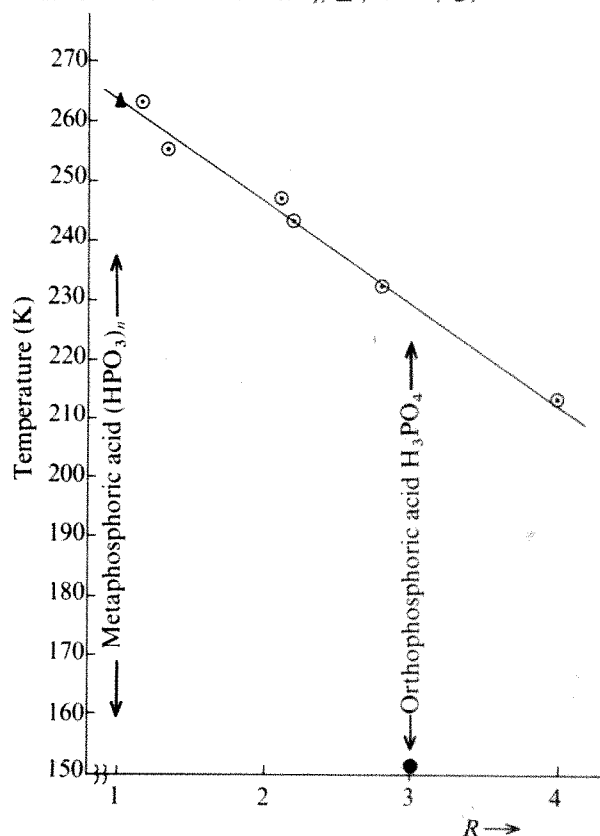
is a linear relationship between T_g and R, when T_g is determined from the inflection point of the proton NMR linewidth temperature curve.

For the metaphosphoric acid, HPO_3 with $R = 1$, T_g has been determined by Eisenberg[11] using a linear variable differential transformer to measure the expansion of the sample with increase of temperature. There is agreement within 1 K between Eisenberg's determination of 263 K and that calculated from the equation given in Fig. 3 with $R = 1$, even though different routes were used for the preparation of the phosphoric acid samples. In spite of the complicated crystallisation behaviour reported for the metaphosphoric acids subject to different thermal treatments[1] it is clear that either dehydration of phosphoric acid at 300 °C or equilibration of a mixture of phosphorus pentoxide with water, yield products with the same glass transition temperature.

Caution is required in the application of the equation (Fig. 3) and it should not be used for values of $R < 1$. There is increasing cross linking as R is reduced below 1, because of the formation of trifunctional units through the elimination of water by condensation of hydroxyl groups on adjacent chains.

O
||
←~~~~O — P — O~~~~~→
|
O
|
←~~~O — P — O~~~~→
||
O
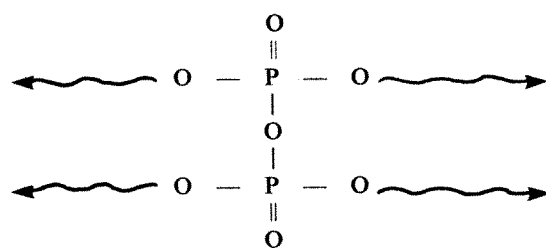
Glass transition temperatures are increased by cross linking but at present it is difficult to test relationships between glass transition temperatures and degree of cross linking. This is because reliable methods for the determination of the concentration of cross links may not be available[4]. Hence because of this increase in cross linking the use of the equation in the region $0 < R < 1$ is inappropriate. For compositions with $R > 4$, experimental determinations of T_g are required before it is possible to check on the applicability of this equation.

When a proton NMR spectrum has a narrow linewidth, $\delta H < 0.25$ gauss, there must be isotropic reorientation of the nuclei at a rate $> 10^4$ Hz, so that $\langle H_{loc} \rangle$ is effectively zero, and the measured line width is determined by the homogeneity of the applied magnetic field and other instrumental factors. When the signal is broad, $\delta H \sim 4-5$ gauss or more, then the protons are in relatively fixed positions. Thus, if a sample has a broad line it must be a solid. For these phosphoric acids the linewidth is broad, up to 5 gauss, even for the sample with $R \sim 4$ at $T < 200$ K. Thus this sample is a glass at temperatures < 200 K. While it is possible for an amorphous material to have molecular motion of various types at $T < T_g$, it is not possible for a sample to have a broad NMR line at $T > T_g$. Therefore, T_g for all these samples must be > 200 K. Kobeko's estimate of 152 K for the glass transition temperature of orthophosphoric acid is far too low.

The use of the definition $T_g = T_i$ (Fig. 2) together with application of our equation, allows estimation of glass transition temperatures in the composition range $1 \leqslant R \leqslant 4$. For orthophosphoric acid, $R = 3$, the predicted glass transition temperature is 229 ± 2 K.

BRYAN ELLIS

Department of Ceramics, Glasses and Polymers,
The University of Sheffield,
Sheffield S10 2TZ, UK

1 Toy, A. D. F., *The Chemistry of Phosphorus,* 482 (Pergamon, Oxford, 1975).
2 Kobeko, P. P., Kuvshinskii, E. V., and Shishkin, N. J., *J. phys. Chem., Mosc.,* 9, 387 (1937).
3 Ellis, B., and McDonald, M. P., *J. non-cryst. Solids,* 1, 186 (1969).
4 Ellis, B., in *Amorphous Materials* (edit. by Douglas, R. W., and Ellis, B.), 357 (Wiley, New York, 1972).
5 Gee, G., *Amorphous Materials* (edit. by Douglas, R. W., and Ellis, B.), 115 (Wiley, New York, 1972).
6 Yates, M. J., thesis, Univ. Sheffield (1973).
7 Smith, R. J., thesis, Univ. Sheffield (1976).
8 Maklakov, A. I., and Pimenov, G. G., *J. polym. Sci., U.S.S.R.,* 15, 123 (1973).
9 McCall, D. W., *Acc. chem. Res.,* 4, 223 (1971).
10 Bondi, A., *Physical Properties of Molecular Crystals, Liquids and Glasses,* 379 (Wiley, New York, 1968).
11 Eisenberg, A., Farb, H., and Cool, L. G., *J. polym. Sci.,* A-2, 4, 855 (1966).

Regenerative failure of double half limbs in *Notophthalmus viridescens*

IT HAS been proposed that positional information in epimorphic fields is specified in terms of polar coordinates[1], where one component of positional information is a value corresponding to position on a circle, and the other to position on a radius. In the newt limb, the circular sequence of positional values lies around the circumference and the radial sequence lies in the proximal–distal axis of the limb. French *et al.*[1] have proposed that various different regulative phenomena in such epimorphic fields as insect legs, insect imaginal disks, and amphibian limbs can be accounted for by two principles of cellular behaviour. The first of these is shortest intercalation, whereby any discontinuities in positional information are resolved by intercalation of missing positional values during growth. The second principle is that distal transformation can only take place from a complete set of positional values in the circular sequence. Unless a complete circle is either present at a site of amputation, or can be generated by intercalation from an incomplete circle, no regeneration (distal transformation) will occur. These principles have been used to account for the number, location, handedness and orientation of supernumerary regenerates produced after grafting experiments in amphibians[2] and insects[3]. The results reported here directly support the idea that a complete circle of positional values in the circular sequence is necessary for limb regeneration in adult newts.

Double half limbs were made in *Notophthalmus viridescens* as shown in Fig. 1. Controls were sham operated limbs in which either an anterior or a posterior half of an upper arm had been removed and replaced in normal orientation. All grafts healed into place rapidly and became fully vascularised by an average of 14 d after operation. The regenerative ability of the controls and the double half limbs was tested by amputating through the grafted region, close to its distal end, usually after 3–4 weeks, but sometimes after 8–9 weeks (eight double half limbs).

All 18 control limbs regenerated normally (Table 1 and Fig 2e), showing that the techniques used do not impair regenerative ability, and that an interval of 3–4 weeks between grafting and amputation is sufficient for restoration of an adequate nerve supply at the amputation plane. When double half limbs were amputated, however, regeneration either failed or was defective. The results were similar for both double anterior and double posterior limbs, and for limbs which were amputated 3–4 weeks and 8–9 weeks after grafting. Of 16 double anterior limbs tested (Table 1), none regenerated normally, and of 18 double posterior limbs,

Table 1 Regenerative ability

Limb type	Total	Defective or no regeneration	Normal regeneration
Control posterior	11	0	11
Control anterior	9	0	9
Double posterior	18	17	1
Double anterior	16	16	0
Half posterior	32	5	27
Half anterior	30	4	26

Fig. 1 Diagram of the grafting operation. *a*, Left and right limbs, cut away distally, to show the major anatomical features. The major muscles of the upper arm are the anconeus, coracobrachialis and humeroantebrachialis. A, Artery; N, nerve; V, vein. To make a double anterior limb, a 1.5-mm long section of the posterior half of a right upper arm was removed by making mid-dorsal and mid-ventral incisions, as well as proximal and distal incisions in the posterior half of the limb, down to the bone. The loosened posterior half was pulled away and discarded. An anterior half limb of matching dimensions from the left upper arm was grafted on to the right limb in place of the discarded posterior half. The graft was oriented normally with respect to its dorso–ventral and proximal–distal axes. Double posterior limbs were made by grafting the posterior half of the right upper arm in place of the anterior half of the left. *b* and *c*, Schematic representations of the operations to form double anterior and posterior limbs. Key positional values in the circular sequence are indicated by numbers from 0 to 12. Operations were performed using chloretone anaesthesia. In all cases, the mid-ventral nerves and artery were left undisturbed on the limb stump. All grafts were stained for 2 min in a 0.05% solution of methylene blue in Holtfreter solution[11] before grafting. Grafts were held in place with small pieces of damp lens paper, and the operated animals were kept cold (10 °C) for 2 d. Between the time that the animals were removed from the cold and the time that their limbs were amputated, they were kept at room temperature (20 °C). Thereafter they were maintained at 25 °C for the remainder of the experiment (2–2.5 months).
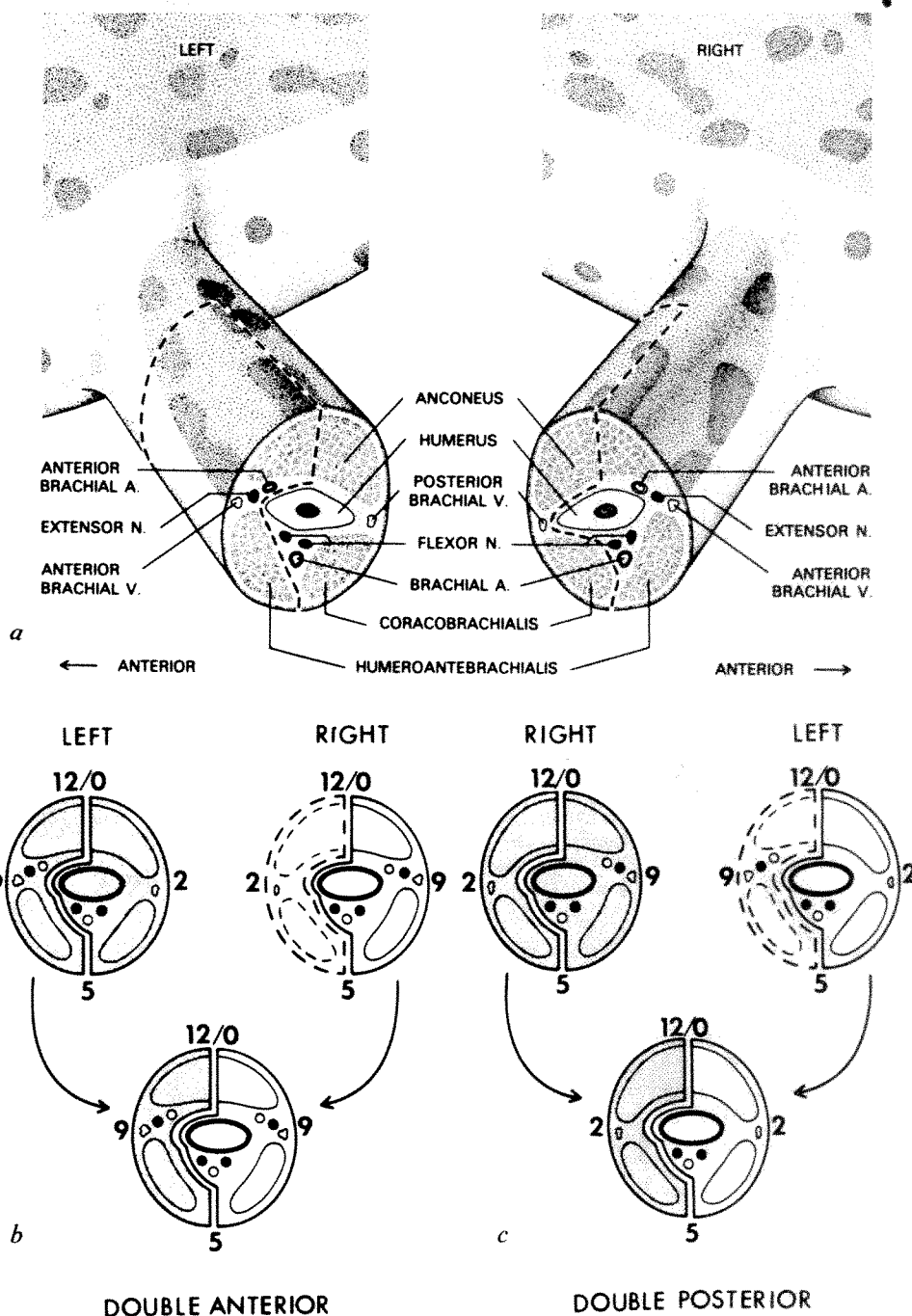
only one formed a normal distal regenerate. The range of skeletal structures formed by the double half limbs 2–2.5 months after amputation varied from little more than a cartilaginous cap on the cut end of the humerus (Fig. 2*c*) or a small separate cartilaginous nodule (Fig. 2*b*) to several small skeletal elements arranged in series (Fig. 2*a* and *d*). These results provide support for the hypothesis of French *et al.*[1] that a complete set of positional values in the circular sequence is necessary for distal transformation in amphibian limbs. The results of preliminary experiments performed on *Ambystoma tigrinum* by Stocum (personal communication) confirm this main result in another species of amphibian.

Previous experiments[4,5] provide further support for the complete circle rule in amphibian limbs. Limbs rendered incapable of regeneration by X irradiation could be caused to regenerate if they were provided with a normally oriented cuff of non-irradiated skin, that is, one containing a complete set of positional values in the circular sequence. It

was also shown, however, that a cuff of skin which was oriented so that only a single 'quality' (that is dorsal, ventral, anterior or posterior) was present at the amputation site could not cause X-irradiated limbs to regenerate. Although the amount of non-irradiated tissue available for regeneration was the same in both instances, regeneration occurred only when this tissue was oriented so as to produce a complete set of positional values in the circular sequence at the amputation site.

Further evidence for the relationship between a complete circle and normal distal transformation comes from the lateral supernumerary regenerates which sometimes formed in the grafting experiments (Fig. 2*d*). In 10 of 18 double posterior limbs and in one of the 16 double anterior limbs (but never in the control limbs), regenerative changes leading to the formation of a complete lateral supernumerary limb were initiated at the proximal graft–host junction when the limb was amputated through the distal region of the graft. A complete, lateral circle of positional values is
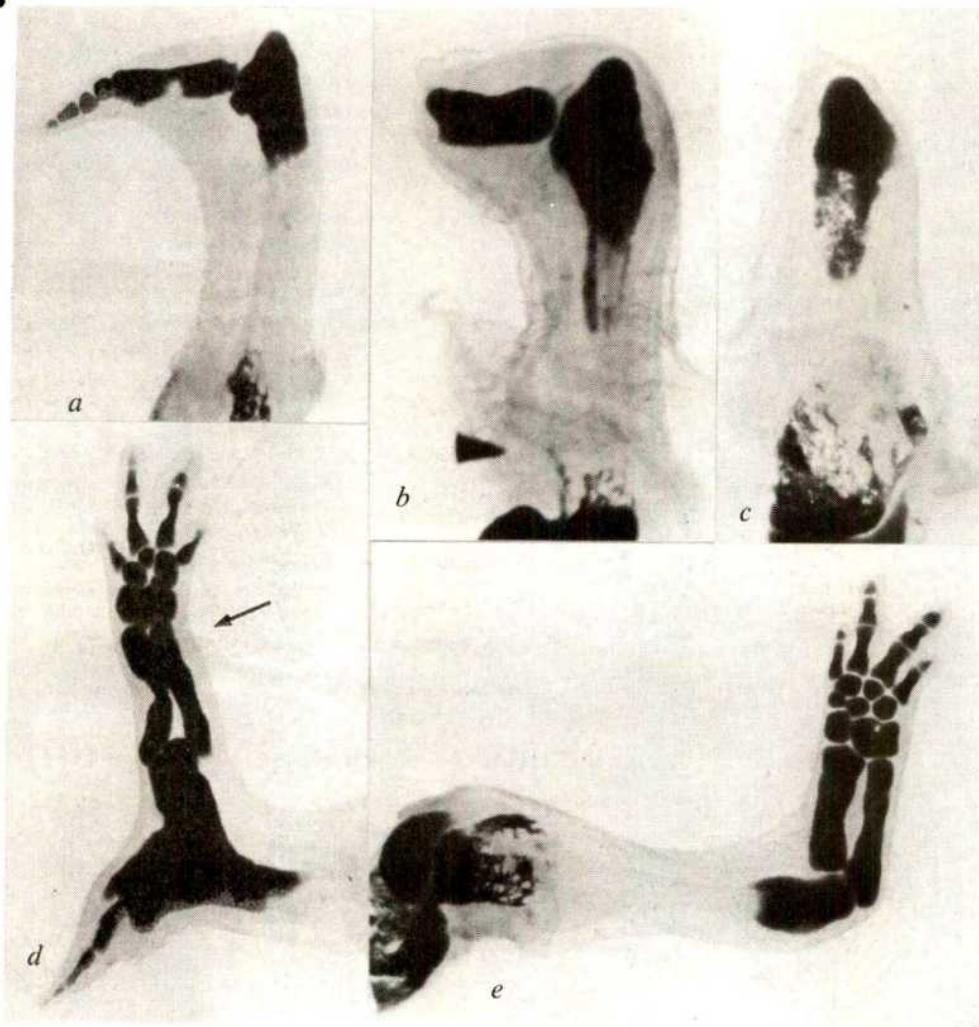
Fig. 2 Skeletal preparations of control and experimental limbs, prepared by the Victoria blue technique[12] 2–2.5 months after amputation. *a* and *b*, Double anterior limbs. *c* and *d*, Double posterior limbs, showing the spectrum of structures produced at the distal amputation site. In (*d*) a lateral supernumerary regenerate (arrow) has developed from the proximal graft–host junction. *e*, A control limb in which the posterior half of the upper arm had been removed and replaced.

created at this position in the experimental limbs (Fig. 3), with half of the values provided by the host and half by the graft. Preliminary evidence from similar grafts made in axolotls indicates that such supernumerary limbs contain cells of both host and graft origin (my unpublished results with Krasner).

The details of the mechanisms involved in distal transformation, and the need for a complete set of values in the circular sequence are unknown. In spite of their failure to complete distal transformation, in these experiments all double half limbs initiated regeneration. Except for one case, however, none proceeded beyond the stage of late bud[6] and the majority were arrested at earlier stages.

A possible alternative interpretation of the failure of double half limbs to regenerate normally is that they have an insufficient number of nerve fibres at the site of amputation. This explanation is untenable for several reasons. First, the major mid-ventral limb nerves were left intact in all cases (Fig. 1). Second, control limbs always regenerated normally. Third, doubling the interval between grafting and amputation from 3–4 weeks to 8–9 weeks, to allow even longer for innervation to occur, did not alter the regenerative performance of double half limbs. Finally, more evidence comes from a multiple grafting experiment in which a double anterior limb is made, left to become vascularised and innervated, then the original host anterior half is replaced by a grafted posterior half, to form a reversed but normal limb. When such limbs are amputated after an appropriate interval, regeneration of a disharmonic limb occurs as expected from the reversed but complete limb stump (my unpublished results). These preliminary results, in addition to the other evidence cited here, indi-

cate that the ability of double half limbs to become adequately innervated is not impaired, and that this alternative interpretation of the failure of double half limbs to regenerate can be rejected.
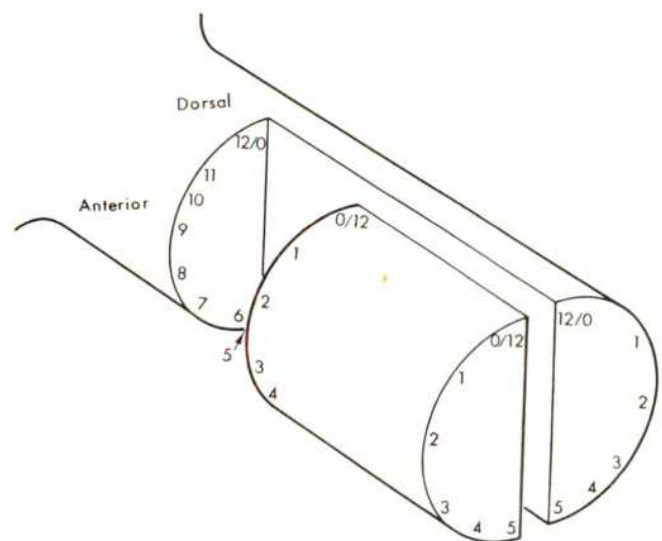
Fig. 3 Diagram of a double posterior limb stump indicating the complete circle of positional values at the proximal graft–host junction. The placement of positional values around the circumference is slightly non-uniform, as suggested by previous results and discussed in ref. 2. The graft is shown to be shorter than it actually is in order that the proximal graft–host junction can be clearly visualised.

In these experiments the regenerative ability of half limbs was also examined (Table 1). In each animal in which one double half limb was produced, the limb from which the graft was taken remained as a half limb, either half anterior or half posterior. These half limbs, as well as half limbs produced on both forelimbs of the same animal were either amputated immediately through the most distal portion of the half limb, or amputated 3–4 weeks later. Most of these limbs developed complete and normal regenerates, thus confirming earlier findings that half limbs can regenerate normally[7,8] or relatively normally[9]. The initial outgrowth did not arise at the most distal cut end of the half limb, but was eccentrically located towards the distal end of the lateral wound edge. During the later stages of regeneration, the regenerate became more symmetrically disposed until by the stage of late digits, little or no asymmetry could be detected. A tentative interpretation of this finding is that the eccentric location of the initial outgrowth is the position on the wound surface at which cells comprising a complete circle of positional values come together during wound healing from the proximal and distal half limb stumps. This possibility is being investigated.

The many similarities between the regulative behaviour of insect and amphibian limbs have led to the hypothesis that the mechanisms of pattern regulation in these distantly related organisms are the same[1]. This idea is strengthened by the similarity of the results reported here to those reported by Bohn[10] with double dorsal and double ventral cockroach legs.

I thank Julie Van Houten for pilot experiments on this project, Warren Fox for assistance and Carol Wilde for art work. This research was supported by a grant from the NIH.

SUSAN V. BRYANT

Department of Developmental
and Cell Biology, and
Center for Pathobiology,
University of California,
Irvine, California 92717

1 French, V., Bryant, P. J., and Bryant, S. V., *Science*, **193**, 969–981 (1976).
2 Bryant, S. V., and Iten, L. E., *Devl Biol.* **50**, 212–234 (1976).
3 French, V., *J. Embryol. exp. Morphol.*, **35**, 267–301 (1976).
4 Carlson, B. M., *Devl Biol.* **39**, 263–285 (1974).
5 Lheureux, E., *Wilhelm Roux's Arch.*, **176**, 303–327 (1975).
6 Iten, L. E., and Bryant, S. V., *Wilhelm Roux's Arch.*, **173**, 263–282 (1973).
7 Weiss, P., *Arch. Entwicklungsmech.*, **107**, 1–53 (1926).
8 Weiss, P., *Principles of Development* (Holt, New York, 1939).
9 Goss, R. J., *J. Morph.*, **100**, 547–563 (1957).
10 Bohn, H., *Wilhelm Roux's Arch.*, **156**, 449–503 (1965).
11 Jacobson, A. G., in *Methods in Developmental Biology* (edit. by Wilt, F. H., and Wessells, N. K.), (Crowell, New York, 1967).
12 Bryant, S. V., and Iten, L. E., *Wilhelm Roux's Arch.*, **174**, 90–101 (1974).

Detection and repair of single-strand breaks in nuclear DNA

DIRECTLY or indirectly phosphodiester bonds in DNA are broken when living cells are irradiated by ionising radiations or ultraviolet light. There are various sophisticated techniques for monitoring radiation damage[1–6]. We describe here how radiation damage in DNA and its repair can be detected simply in the white cells of human blood. The method is very sensitive and should prove useful in screening populations for abnormal repair mechanisms. As there is also great interest in methods for detecting environmental agents that damage DNA[2,7], we have applied the method to detect the damage caused by mitomycin C.

We have described our technique for detecting the damage caused to the nuclear DNA of HeLa cells by very low doses of γ rays and also for monitoring the repair of this damage[8,9]. When cells are lysed in the presence of non-ionic detergents and high salt concentrations, structures resembling nuclei—called nucleoids—are released. These nucleoids contain nearly all nuclear RNA and DNA, but are depleted of nuclear proteins[10]. Their DNA is supercoiled and compact so that the nucleoids sediment more rapidly in sucrose gradients than their γ-irradiated counterparts which contain extended DNA with single-strand breaks or nicks. Repair of the breaks restores the DNA to its original conformation and re-establishes the normal sedimentation rate. We monitor DNA integrity by measuring the distance sedimented by nucleoids in sucrose gradients.

Nucleoids are released immediately on addition of white blood cells to a lysis mixture containing the non-ionic detergent Triton X-100 and 1.95 M NaCl. The effects of the intercalating dye, ethidium bromide (EB), on the sedimentation of nucleoids derived from white blood cells was studied by spinning the nucleoids through sucrose gradients containing 1.95 M NaCl and different concentrations of the dye (Fig. 1). The distance travelled by the nucleoids is expressed as a ratio relative to that travelled by unirradiated nucleoids sedimenting in the same conditions but in the
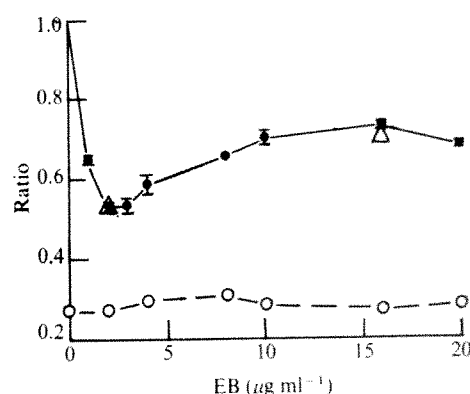
Fig. 1 Effect of EB on the sedimentation of nucleoids from human white blood cells. The distance sedimented by nucleoids in gradients containing different concentrations of EB is expressed as a ratio relative to that of unirradiated nucleoids sedimenting in the absence of EB. ●, Unirradiated nucleoids; ○, nucleoids γ irradiated (960 rad, 120 rad min^{-1}) after addition of cells to the lysis mixture[8]; △, unirradiated nucleoids from a patient with xeroderma pigmentosum. Error bars give the standard error of the mean. Human white cells were obtained from peripheral blood[17] and cultured in RPMI 1640 medium supplemented with 20% foetal calf serum (Flow Laboratories). Between 1 and 5% of the cells prepared in this way are red cells, the remainder are almost all lymphocytes. The results obtained with blood from four normal males and two females were pooled. In some cases white cells stored in medium on ice for 24 h were used; the sedimentation properties of nucleoids prepared from such cells were similar to those obtained from freshly isolated cells. Samples of 50 μl of a suspension containing between 2 and 5×10^5 cells in phosphate-buffered saline (PBS) were layered on 150 μl of a lysis mixture floating on top of "isokinetic" sucrose gradients[8,10]. Gradients (15–30% sucrose, 4.6 ml, pH 8.0) contained 1.95 M sodium chloride, 0.01 M Tris, 0.001 M EDTA in addition to various concentrations of EB. The lysis mixture contained sodium chloride, EDTA, Tris and Triton X-100 in amounts which, on addition of 1 volume of PBS containing cells to 3 volumes of the mixture, gave final concentrations of the constituents of 1.95 M, 0.1 M, 2 mM and 0.5% respectively. (For the purpose of calculating the final concentration of sodium chloride, the contribution of the PBS is neglected). Fifteen minutes after the addition of the cells to the lysis mixture, gradients were spun at 30,000 r.p.m. for 25 min at 20 °C in the SW50.1 rotor in a Beckman L2-65b ultracentrifuge. After the gradients had been spun, the position of the nucleoids in the gradient was determined by their absorbance at 254 nm. One gradient of the six spun in the rotor served as a reference; the distance travelled by nucleoids in other tubes is expressed relative to the distance sedimented by nucleoids in the reference tube[8].

absence of EB. As the concentration of EB in the gradient is increased, the distance travelled by nucleoids falls to a minimum and then rises again. γ Irradiation reduces the sedimentation rate of the nucleoids and abolishes the biphasic response to EB. The shape of a similar curve obtained with HeLa nucleoids is discussed elsewhere[9]. As the sedimentation rate of supercoiled DNA varies in this characteristic biphasic manner[11], we conclude that nucleoid DNA is supercoiled. Irradiation introduces single-strand breaks into the DNA so that supercoils are lost.
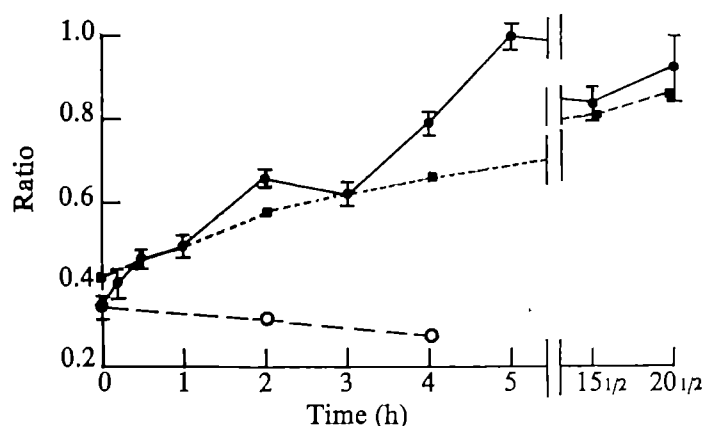
Fig. 2 Effect of incubation of white blood cells after γ irradiation on nucleoid sedimentation. The distance sedimented by nucleoids from irradiated cells in gradients lacking EB is expressed as a ratio relative to that of nucleoids from unirradiated cells which had been treated similarly. White cells (2 × 10⁴ per ml in medium) were irradiated (960 rad)[9], and incubated for different times at 37 or 4 °C. They were then added to 10 volumes of ice-cold PBS, pelleted and resuspended in ice-cold PBS. Samples of 50 µl were applied to gradients which were then spun and analysed as described in the legend to Fig. 1. Error bars give the standard error of the mean. Normal white blood cells, irradiated and incubated at 37 °C (●) or 4 °C (○). White cells from a patient with xeroderma pigmentosum, irradiated and incubated at 37 °C (■).

Repair of the damage caused by γ irradiation was studied by irradiating white blood cells (960 rad) and then incubating them for different periods before applying them to gradients lacking EB and then measuring the sedimentation rate of the nucleoids. (A dose of 960 rad reduces the cloning efficiency of HeLa cells by 99% (ref. 8). Incubation at 37 °C after irradiation, but not at 4 °C, increased the sedimentation rate of the nucleoids (Fig. 2). After 5 h, the sedimentation rate had been restored almost to that of un-irradiated nucleoids, and the nucleoids from the irradiated cells had regained their biphasic response to EB (Table 1).

These techniques have been extended to include measurements on the effects of ultraviolet light. Repair of damage induced by ultraviolet light involves cutting one strand of the DNA duplex ("incision"), removal of the principal photoproduct, the thymine dimer ("excision"), synthesis of DNA complementary to the unaffected strand and "ligation" of the final phosphodiester bond to restore the intact duplex[9]. HeLa cells were irradiated (25 erg mm⁻²) and incubated for different times at 37 °C before the sedimentation rate of nucleoids isolated from them was measured in gradients lacking EB (Fig. 3a). (A dose of 25 erg mm⁻² reduces the cloning efficiency of HeLa cells by 24% (our unpublished results).) Irradiation without subsequent incubation at 37 °C reduces the rate of sedimentation. This probably results from the single-strand breaks induced by the highly active incision enzymes of HeLa cells during the time taken to concentrate the cells before they are lysed. (Although the formation of a thymine dimer in DNA unwinds the double helix by 5–6° (ref. 12), the sedimentation rate is unlikely to be reduced by such unwinding

in the absence of strand scission because irradiation of isolated nucleoids with high doses of ultraviolet light does not alter the sedimentation rate (our unpublished results).) Incubation of the cells at 37 °C for 5 min after irradiation reduced the rate of sedimentation of the nucleoids further, whereas incubation periods greater than 20 min increased the distance sedimented until it became indistinguishable from that of the unirradiated controls. We conclude that after irradiation, incision initially leads to a partial loss of supercoils in nucleoid DNA, but then ligation restores supercoiling (Fig. 3a).

A measure of the damage caused by different doses of ultraviolet light is given by the distance sedimented in the absence of EB by nucleoids after the irradiated cells had been incubated for 5 min at 37 °C to maximise the number of strand breaks (Fig. 4). As radiation dose increased, the distance sedimented decreased progressively.

The damage caused by ultraviolet light is repaired much more slowly in white blood cells than in HeLa cells. Strand breakage and loss of supercoiling were maximal 2 h after irradiation. The normal supercoiled configuration was not fully recovered after 15.5 h (Fig. 3b and Table 1). There was
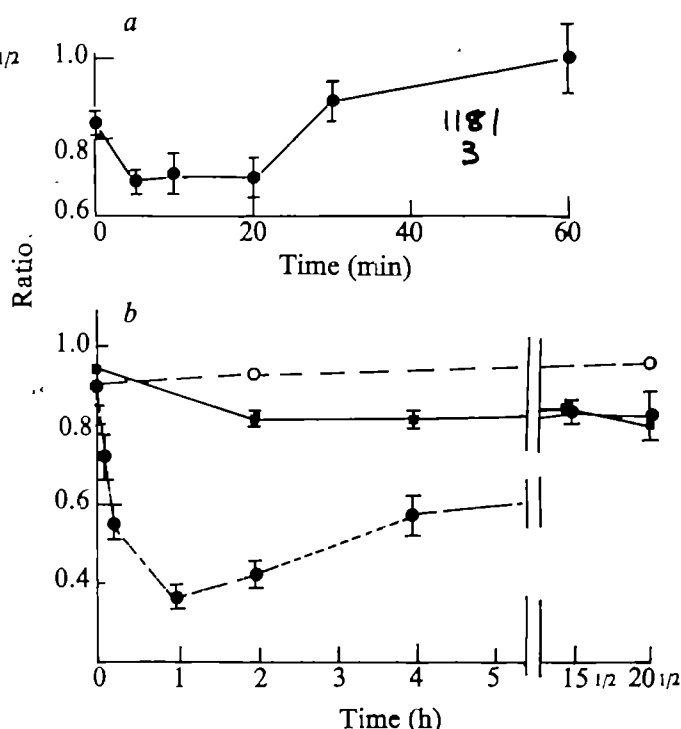
Fig. 3 Effect on nucleoid sedimentation of incubating (a) HeLa cells and (b) human white blood cells at 37 °C after irradiation with ultraviolet light. a, Human HeLa cells were grown in suspension[9]. 0 5 ml of cells (2 × 10⁴ per ml) in ice-cold medium in a Petri dish were irradiated (25 erg mm⁻¹, 5 erg mm⁻¹ s⁻¹) with ultraviolet light while the Petri dish was shaken vigorously. The source of ultraviolet light was a Sylvania germicidal tube (G 15T8). Doses were measured with an ultraviolet meter[14]. After irradiation, cells were mixed with 4.5 ml of medium at 37 °C and incubated in the dark at 37 °C. At the end of incubation the cells were mixed with 9 volumes of ice-cold PBS, pelleted and resuspended in ice-cold PBS. Samples of 50 µl of the cell suspension were applied to sucrose gradients lacking EB. Gradients were spun at 5,000 r.p.m. for 50 min and analysed as described in the legend to Fig. 1. The distance sedimented by nucleoids derived from irradiated cells is expressed as a ratio relative to the distance sedimented by nucleoids derived from cells treated similarly but which had not been irradiated. Error bars give the standard error of the mean. b, The experiments with white blood cells were similar to those with HeLa cells except that media and spinning conditions were different (see legend to Fig. 1). ●, Normal white blood cells, irradiated and incubated at 37 °C; ○, normal white blood cells, irradiated and incubated at 4 °C; ■, white blood cells from a patient with xeroderma pigmentosum, irradiated and incubated at 37 °C.

Table 1 Effect of different treatments on supercoiling in the nucleoids from white blood cells

Treatment	Relative distance sedimented by nucleoids in the presence of EB			
	$0 \, \mu g \, ml^{-1}$	$2 \, \mu g \, ml^{-1}$	$4 \, \mu g \, ml^{-1}$	$16 \, \mu g \, ml^{-1}$
Untreated	1.0	0.53	0.59	0.73
4-h incubation at 37 °C after γ irradiation (960 rad)	0.82	0.53	0.66	0.65
2-h incubation at 37 °C after ultraviolet irradiation (25 erg mm^{-2})	0.39	0.39	—	0.35
15.5-h incubation at 37 °C after ultraviolet irradiation (25 erg mm^{-2})	0.84	0.58	—	0.68

Nucleoids obtained from cells treated in different ways were spun in gradients containing different concentrations of EB. The distance sedimented by nucleoids from treated cells is expressed as a ratio relative to that of nucleoids from untreated cells sedimenting in the absence of ethidium. The conditions used for the untreated cells, and cells irradiated with γ rays or ultraviolet light are described in the legends to Figs 1, 2 and 3, respectively.

little change in the sedimentation rate of nucleoids when the irradiated cells were incubated at 4 °C (Fig. 3*b*).

Cells from patients with the autosomal recessive disease, xeroderma pigmentosum can repair the damage caused by ionising radiation but cannot effectively repair DNA damaged by ultraviolet light. Some forms of this disease are thought to be characterised by deficiency in "incision" activity[5]. Nucleoids prepared from the white cells of our patient sedimented in gradients containing EB exactly as the nucleoids from normal cells (Fig. 1). In the absence of EB the sedimentation behaviour of nucleoids prepared from γ-irradiated cells was also indistinguishable from that of nucleoids prepared from normal cells that had been irradiated (Fig. 2): the patient's cells could thus repair γ-ray damage. On the other hand, irradiation with ultraviolet light affected normal cells and those of the patient differently (Fig. 3*b*). When the patient's cells were irradiated with ultraviolet light and incubated at 37 °C the sedimentation rate of the nucleoids in the absence of EB was not reduced as much as that of controls : the patient's cells were unable to "incise" normally and the nucleoids retained supercoils and sedimented rapidly.

Mitomycin C is a carcinogen[7]. After activation by cellular enzymes it cross links DNA strands without breaking them. Strands are broken only when the damaged region is removed by repair enzymes[13]. We investigated the effects
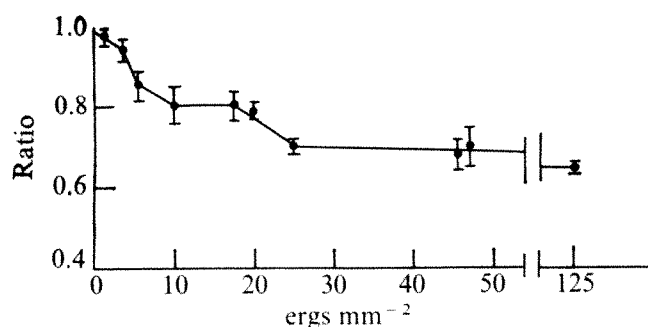
Fig. 4 Effect of different doses of ultraviolet radiation on the sedimentation of HeLa nucleoids. HeLa cells were irradiated for 5 s with different doses of ultraviolet light, incubated for 5 min at 37 °C, applied to sucrose gradients lacking EB and the gradients were spun and analysed (see legend to Fig. 3). The distance sedimented by nucleoids derived from irradiated cells is expressed as a ratio relative to the distance sedimented by nucleoids derived from cells which had been treated similarly but which had not been irradiated. Error bars, s.e.m.

of mitomycin C on the integrity of DNA by incubating white cells at 37 °C for 2 h in different concentrations of the drug, lysing the cells, and determining the rate of sedimentation of the nucleoids in the absence of EB (Fig. 5). Low concentrations of the drug had no effect on the distance sedimented, higher concentrations reduced it. At the high concentrations, strand breakage caused loss of supercoiling and slow sedimentation. The effects of lower concentrations can be detected by incubating the cells with the drug for longer periods (our unpublished results). If, after a 2-h exposure to the drug, the cells were washed and incubated

in its absence at 37 °C for 15.5 h, faster sedimentation was restored, repair of the damage restored normal supercoiling.

We have made nucleoids containing superhelical DNA from a wide range of cells (fibroblasts, lymphoblasts, erythroblasts, teratocarcinoma and epithelial cells of men, mice, birds, amphibians and insects (ref. 9 and our unpublished observations)), so that these methods for detecting single-strand breaks in DNA should be applicable to most cells.
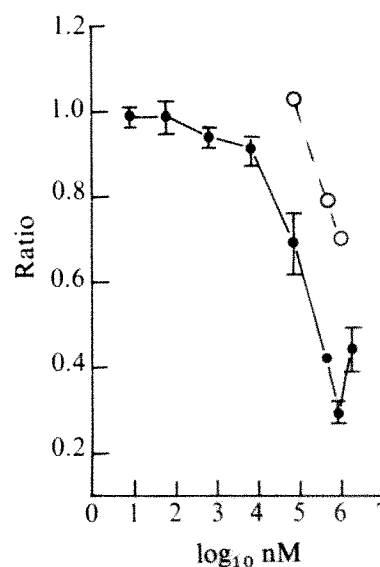
Fig. 5 Effect of mitomycin C treatment of white cells on the sedimentation of nucleoids in gradients lacking EB. White cells were incubated for 2 h at 37 °C (see legend to Fig. 1) in the presence of different concentrations of mitomycin C (Boehringer). Some cells were then collected and applied to gradients immediately and the gradients were spun and analysed (see legend to Fig. 1). Other cells were washed twice with medium, resuspended at 2×10^6 per ml in medium lacking mitomycin C and incubated a further 15.5 h at 37 °C. They were then collected and treated as the others.. The distance sedimented by nucleoids derived from cells treated with mitomycin C is expressed as a ratio relative to the distance sedimented by those derived from untreated cells. Error bars, s.e.m. ●, Cells incubated for 2 h in the presence of mitomycin C. ○, Cells incubated for 2 h in the presence of mitomycin C, washed and then incubated for a further 15.5 h at 37 °C.

Using white cells obtained from human blood taken in the morning, we can assess by the evening whether repair of damaged DNA in the cells is normal. While doses of 960 rad (γ rays) or 25 erg mm^{-2} (ultraviolet light) produce the large effects described in these experiments, much lower doses (that is, 12 rad and 5 erg mm^{-2}) can be detected. The method should therefore be particularly well suited to the rapid screening of patients for deficiencies in the mechanisms that repair radiation damage.

Many carcinogens and mutagens break phosphodiester bonds in cellular DNA either directly or indirectly after enzyme action. Breakage, which can be demonstrated by

conventional physical techniques[14] probably increases repair synthesis[15] and recombination[3,16]. Agents that break DNA can be detected in nucleoids with the sensitivity, rapidity and economy usually associated with the use of bacteria[7]. There may well be advantages in using human cells in any screening method for detecting potential human carcinogens.

We thank Dr F. Giannelli for help and for white blood cells from a patient with xeroderma pigmentosum, and Professor H. Harris for support and encouragement. We thank the SRC and the CRC for support.

P. R. COOK
I. A. BRAZELL

Sir William Dunn School of Pathology,
South Parks Road,
University of Oxford,
Oxford, UK

1 Kao, F. T., and Puck, T. T., *J. cell comp. Physiol.*, 74, 245–258 (1969).
2 Perry, P., and Evans, H. J., *Nature*, 258, 121–125 (1975).
3 McGrath, R. A., and Williams, R. W., *Nature*, 212, 534–535 (1966).
4 Fornace, A. J., Kohn, K. W., and Kann, H. E., *Proc. natn. Acad. Sci. U.S.A*, 73, 39–43 (1976).
5 Cleaver, J. E., *Adv. radiat. Biol.*, 4, 1–75 (1974).
6 Grossman, L., *Adv radiat Biol.*, 4, 77–129 (1974).
7 McCann, J., and Ames, B. N., *Proc. natn. Acad. Sci U S.A*, 73, 950–954 (1976).
8 Cook, P. R., and Brazell, I. A., *J. Cell Sci.*, 19, 261–279 (1975).
9 Cook, P. R., and Brazell, I. A., *J Cell Sci* (in the press)
10 Cook, P. R., Brazell, I. A., and Jost, E., *J. Cell Sci.* (in the press).
11 Bauer, W., and Vinograd, J, *Prog. molec subcell. Biol*, 2, 181–215 (1971).
12 Denhardt, D. T., and Kato, A. C., *J. molec. Biol* 77, 479–494 (1973).
13 Szybalski, W., and Iyer, V. N., in *Antibiotics*, 1 (edit. by Gottlieb, D., and Shaw, P. D.), 211–245 (Springer, Berlin, 1967).
14 Strauss, B S., in *Chemical Mutagens, Principles and Methods for their Detection*, 1 (edit. by Hollaender, A.), 145–174 (Plenum, New York, 1971).
15 Stich, H. F, in *Molecular Mechanisms for Repair of DNA*, Part B (edit. by Hanawalt, P. C., and Setlow, R B.), 773–784 (Plenum, New York, 1975).
16 Jacob, F., and Wollman, E. L., *Ann Inst. Pasteur*, 88, 724–749 (1955).
17 Thorsby, E., and Bratlie, A., in *Histocompatibility Testing* (edit. by Terasaki, P. I), 655 (Munksgaard, Copenhagen, 1970).
18 Latarjet, R., Morenne, P, and Berger, R, *Ann. Inst. Pasteur*, 85, 174–184 (1953).

Evidence of a thymic abnormality in murine muscular dystrophy

THE involvement of the thymus in muscular, neuromuscular and neurological disorders is well documented[1–4]. In spite of the extensive study made of the thymus in myasthenia gravis, which, as a neuromuscular disorder with clinical symptoms of muscular weakness, shares some similarities with muscular dystrophy, the state and function of the thymus in muscular dystrophy has been ignored. The poor health and susceptibility to infection of dystrophic mice have frequently been observed. These observations and the similarities between dystrophic wasting and that seen in mice thymectomised pre- or neonatally[4], led us to the conclusion that an investigation into the status of the thymus in murine muscular dystrophy was warranted. We have found that the thymuses of dystrophic mice exhibit changes in the age-dependent variation in thymic weight, together with aberrations in cellular morphology indicative of altered secretory activity.

Male and female dystrophic mice of the Bar Harbor 129/ReJ strain were killed by cervical dislocation at various ages, their thymuses, spleens and elbow lymph nodes removed, dissected free of fat and connective tissue, weighed and fixed in 10% formol-saline or methanol. The thymuses, spleens and lymph nodes from clinically normal littermates, and also from a non-dystrophic strain of mice (CBA), were used for comparison. The organs were paraffin embedded and subjected to routine haematoxylin–eosin staining procedures. Unna–Pappenheim staining was also used[5]. The histology of random single sections of these organs was then compared.

For an initial indication of thymic status, the thymus weights of normal and dystrophic mice were compared. It is not possible to compare thymic indices in these animals because of the selective wasting of hind limb muscle in the dystrophic mouse. Generally, mouse thymuses have

been found to decline in weight from 4 to 12 weeks of age; the thymuses of the normal littermates followed this trend. In dystrophic mice, the thymic weights were at first much lower, but then increased with age, reaching a maximum at approximately 9 weeks of age, at a weight comparable to that of the normal mouse at 4–6 weeks of age. The thymic weights of the dystrophic mice then slowly declined to a level approaching that of their normal littermates at 12 weeks of age (Fig. 1). In all strains of mice so far studied, there has been to our knowledge, no report of a similar weight gain to 9 weeks of age, the predominant pattern being an increase in weight to 4 weeks of age,
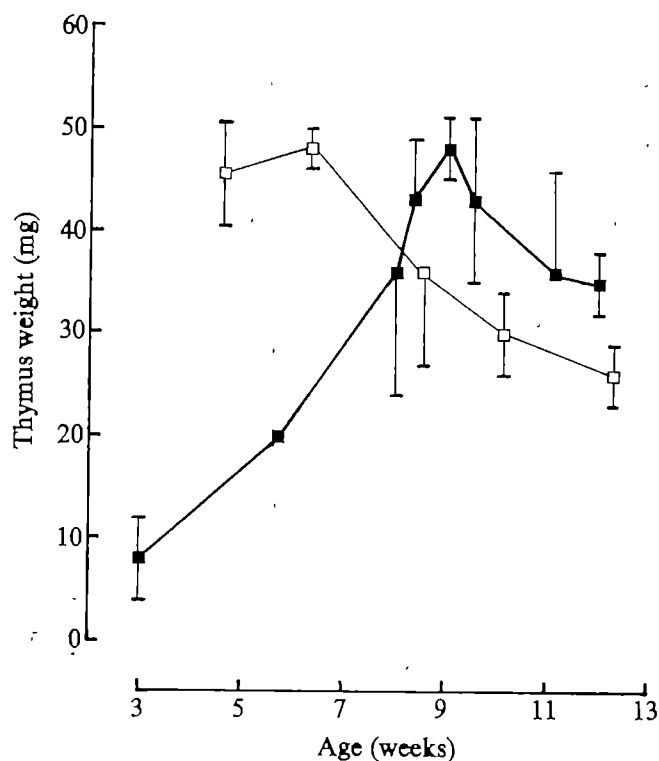
Fig. 1 Thymus weight as a function of age in normal and dystrophic mice. Thymus weight (mg)±s.d. Normal mice, 4–7 animals per group. Dystrophic mice, 2–6 animals per group.

followed by a gradual, but definite, decrease in weight which continues throughout the life span of the animal[7].

This anomaly led us to consider the histology of the thymus of dystrophic mice. As shown in Fig. 2, this also differed dramatically from the normal. While the normal littermates showed no abnormalities of basic thymic morphology, a number of the dystrophic animals studied exhibited remarkable alterations of thymic cellular composition. In the main, this consisted of a depletion, and in some cases, total absence of lymphocytes from lobules of the thymus, these being replaced by epithelial cells which showed very marked signs of secretory activity[4]. In extreme cases, the epithelial cells had formed themselves into classical, though duct-less acini, surrounding a material which did not stain in the conditions used. These drastic changes in thymic composition were found in young (3–8 weeks of age) dystrophic mice which were severely affected by the disease, as evidenced by the degree of disability. In older animals, and those less severely affected, the thymus showed less dramatic changes. These consisted of an apparent relative increase in medullary area over cortical area and a loss of distinction of the cortico–medullary junction. The change in the ratio of medulla to cortex could be due to an increase in medullary tissue or a decrease in the number of lymphocytes populating the
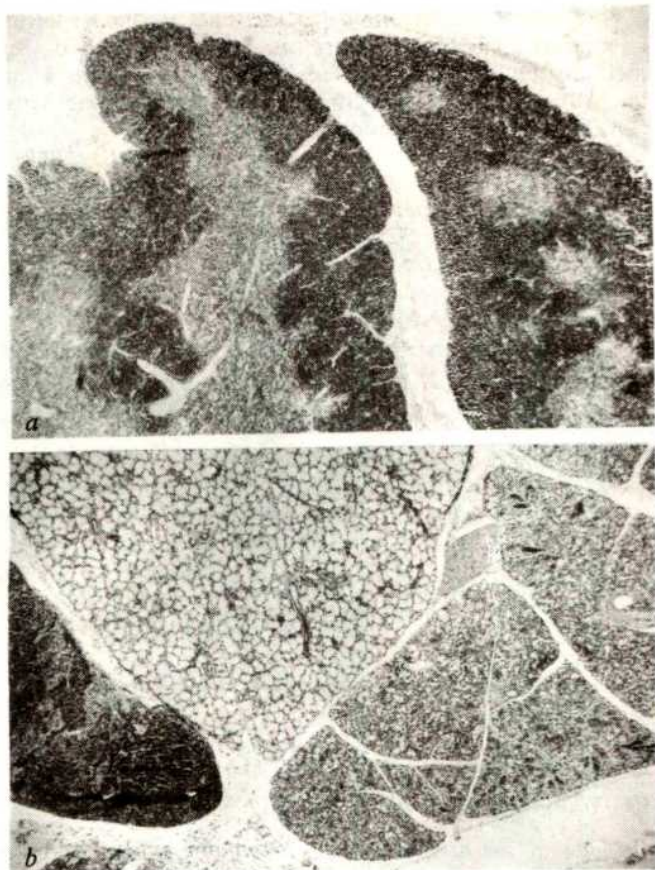
Fig. 2 Histology of thymuses from normal and dystrophic mice. *a*, Thymus from clinically normal littermate. The cortico–medullary junction is distinct; few epithelial cells can be detected in the cortex which is tightly packed with lymphocytes. Haemotoxylin and eosin, ×10.4. *b*, Thymus from severely affected dystrophic mouse. The three lobules shown were enclosed by a continuous capsule of thymic epithelium. The lobule on the left most closely resembles the normal state, but even here, patches of lighter staining cells can be seen in the outer cortex. The lobule on the right represents an intermediary stage in which the loss of definition of the cortico–medullary junction, the depletion of the lymphocytes present in the lobule, and the apparent increase in the number of epithelial cells, was the earliest and most consistent observation. The centre lobule shows the most extreme degree of alteration, having no resemblance to the normal thymus. The nuclei of what are presumed to be epithelial cells because of the presence of cytoplasm around the nuclei form a rim around an accumulation of poorly staining material, the chemical nature of which is unknown. Haemotoxylin and eosin, ×10.4.

cortex. As the packing of lymphocytes in the cortex was noticeably less dense, the latter explanation seems more probable. The cortico–medullary junction, which is normally quite distinct, was indefinite in dystrophic mice, the change from medulla to cortex taking place over a much larger area than is seen in the normal littermates. In some thymuses from dystrophic mice an infiltration of epithelial cells from the outer rim of the cortex, into the cortex, was also seen. The main changes seen in the thymuses in this study are summarised in Table 1.

As all these findings indicated a loss or lack of lymphocytes in the thymus of dystrophic mice, the spleen and lymph nodes were investigated as possible recruiting areas for the missing lymphocytes. Dystrophic spleen weights were found to differ from those of normal mice, the spleens of dystrophic mice being consistently smaller at all ages than those of their normal littermates (Fig. 3). It was also noted that while spleen weights of normal mice generally exhibited a certain variability, reflecting the ability of the

individual animals to respond to infectious agents and other stimulants of the immune system (splenomegaly), the spleen weights of the dystrophic mice were relatively constant. Investigation of splenic histology did not reveal any gross differences between normal and dystrophic mice. It was, however, noted that no germinal centres, as detected by Unna–Pappenheim staining, were seen in the spleens of dystrophic mice, although they were often found in the spleens of normal animals. The lymph nodes of dystrophic animals seemed to contain fewer lymphocytes and relatively more endothelial cells than the lymph nodes of normal animals.

These observations and the peculiarities of thymic weight and thymic histology all indicate a possible defect in function of the immune system in dystrophic mice. The altered morphology of the thymuses of dystrophic mice indicate an abnormal secretory activity of the epithelial cells of the thymus. These cells have been implicated as the site of synthesis and secretion of thymus hormones in normal mice[8,9]. A wide variety of actions have been ascribed to thymic hormones; notably, these include both potentiating and inhibitory actions on all immune functions, the capability of interacting with a variety of other endocrine hormones, and hence actions on growth, muscle regeneration and the capacity of the animal to respond to disease[10]. If, as the evidence suggests, there is a defect somewhere in the system of thymic hormones and/or their actions, this
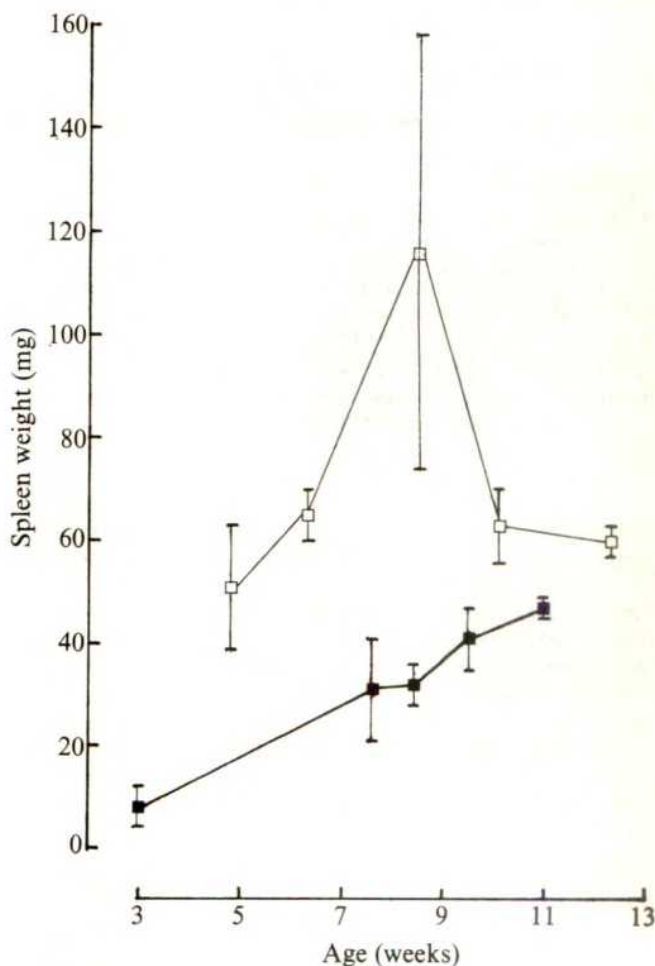
Fig. 3 Spleen weight as a function of age in normal and dystrophic mice. Spleen weight (mg)±s.d. Normal mice, 4–7 animals per group. Dystrophic mice, 4 animals per group. The large variation in spleen weight of normal mice observed at 8.5 weeks of age is interpreted as a reflection of the capacity of these mice to undergo splenomegaly (see text). □, normal; ■, dystrophic.

could explain a well known, but generally ignored, characteristic of murine muscular dystrophy. The disease has a random clinical onset and remarkably variable progression, some animals dying at 3 weeks of age, others surviving to 6 months. If this disease has an hormonal basis, then individual animals would be expected to have different capacities to compensate for an hormonal excess or insufficiency, inactivity or hyper-reactivity, by alteration of other hormones with which the affected hormone interacts. The intricacies of such interaction and compensation between endocrine hormones, and the variation between individuals has been reviewed[11]. Furthermore, the finding that the thymus weights of dystrophic mice follow an age-pattern different from that of normal mice, regardless of severity of disease, or extent of thymic abnormality, suggests that the thymus is abnormal in all dystrophic mice. The effect and extent of this abnormality may be defined by a more complex system, which is individually determined. In this connection, it is worth noting that, as remarked previously, gross signs of epithelial secretion in the thymus were detected only in those dystrophic animals severely affected early in life. In particular, the thymuses of two 3-week-old mice which were close to death from the dystrophic process, were totally epithelial in nature and with a marked degree of secretory activity evident.

Table 1 Histological features of thymuses from normal and dystrophic mice

| | | | Histological features | |
	No. of mice	% normal	% without distinct cortico–medullary junction	% showing signs of epithelial secretion
+/?	40	90	10	5
dy/dy	45	4	96	29
CBA	10	100	0	0

The percentage of abnormalities seen in clinically normal animals (+/?) may be explained by the uncertainty inherent in selecting as normal all animals which do not show overt clinical symptoms of dystrophy. It is possible that in doing this, animals which are dystrophic but have not yet developed the clinical symptons, may be classified as normal. It is for this reason that the CBA mice were included in this study. Further investigations are being made on the thymus of the homozygous normal Bar Harbor 129/ReJ strain of mice.

The effect of the thymus on immune mechanisms is most marked in mice in neonatal and early postnatal life (up to 4 weeks of age); thereafter the influence of the thymus on immunological capacity declines. Also, the signs of epithelial secretion noted in the thymuses of dystrophic mice at various ages after 2 weeks of age, are similar to those found by Clark in normal mice from birth up to an age of 2 weeks[6].

The alteration of thymic weights have been independently confirmed by D. F. Horrobin and R. A. Karmali (see following paper), who found alterations in the responses of thymuses from dystrophic animals to prolactin, and to the suppression of prolactin secretion.

On the basis of these observations, it is postulated that the thymus in murine muscular dystrophy is functionally altered, and that this alteration probably results in the abnormal secretion of one or more thymic hormones. Whether this altered secretion is due to a malfunction within the thymus itself or to a disorder of a system interacting with the thymus, is not known. This and many other possible avenues are yet to be investigated. Further investigation of the thymic abnormality of dystrophic mice, including periodic acid–Schiff staining and ultrastructural studies, are currently underway.

T. A. DE KRETSER
B. G. LIVETT

*Department of Biochemistry,
Monash University, Clayton,
Victoria, Australia 3168*

1 Yunis, E. J., Stutman, O., and Good, R. A., *Ann. NY Acad. Sci.*, **183**, 205–220 (1971).
2 Field, E. J., and Caspary, E. A., *J. Neurol. Neurosurg. Psychiat.*, **36**, 604–606 (1973).
3 Minderhoud, J. M., and Teelker, A. W., *Clin. Neurol. Neurosurg.*, **77**, 225–233 (1974).
4 Minderhoud, J. M., Teelker, A. W., Leenstra-Borsje, H., and Oosten, J. K., *Clin. Neurol. Neurosurg.*, **77**, 234–240 (1974).
5 Miller, J. F. A. P., in *Thymus in Immunobiology* (edit. by Good, R. A., and Gabrielsen, A. E.), 436 (Hoeber–Harper, New York, 1964).
6 *Manual Histopathological Staining Methods* (edit. by Putt, F. A.), (Wiley, New York, 1972).
7 Belyaev, D. K., and Gruntenko, E. V., *Nature*, **237**, 401–402 (1972).
8 Arnesen, K., *Acta Path. Micro. scand.*, **43**, 339–350 (1958).
9 Clarke, S. L., in *The Thymus. Experimental and Clinical Studies* (edit. by Wolstenholme, G. E. W., and Porter, R.), 3–38 (Churchill, London, 1966).
10 Luckey, T. D., (ed.), *Thymic Hormones* (University Park Press, Baltimore, 1973).
11 Horrobin, D. F., *Prolactin: Physiology and Clinical Significance* (Medical and Technical Publ. Co., 1973).

Abnormalities of thymus growth in dystrophic mice

IN spite of the attention paid to the relationship between myasthenia gravis and the thymus, little work has been done on the thymus in muscular dystrophy. Walker[1] found that mouse dystrophy was not alleviated by neonatal thymectomy and Matheson[2] provided evidence against the idea that muscular dystrophy might be a lymphocyte-mediated autoimmune disease. It therefore seems unlikely that overactivity of lymphoid function of the thymus has any role in muscular dystrophy: the possible role of a thymic deficiency has not been considered. Our interest was aroused when we noticed that in certain superficial ways dystrophic mice of the Bar Harbor 129 strain[3–5] appear similar to mice which have been neonatally thymectomised. Their growth is stunted to a very variable degree, their coat quality is poor and they are liable to die suddenly for no very obvious reason. We now present evidence for an abnormality of thymus growth in these animals.

Initially we simply compared body and thymus weights in untreated 4-week and 12-week-old mice of the Swiss and Bar Harbor 129/ReJ strains. At 4 weeks, 1 week after weaning, it is usually possible to tell by observation whether a 129 animal is dystrophic (dy/dy) or phenotypically normal (+/+ or +/dy). The presence or absence of dystrophy was verified by muscle histology. The results are shown in Table 1. There were approximately equal numbers of each sex in each group and no sex differences were noted. At both 4 and 12 weeks the Swiss mice were significantly larger than phenotypically normal 129 mice (P<0.01)—which in turn were larger than the dystrophic mice. At 4 weeks there were no differences in absolute thymus weights between the Swiss and normal 129 mice but in the dystrophic mice the glands were very much smaller (P<0.001). At 12 weeks the normal reduction in thymus weight was well advanced in the Swiss mice but less so in the normal 129s. In the dystrophic animals thymus weight was actually greater than at 4 weeks (P<0.01). Because of the known effects of adrenal steroids on the thymus, in some of the dystrophic and phenotypically normal 4-week-old animals, we measured adrenal weights as well. In the normals mean weight of the adrenals was 3.4±0.5 (s.e.m.) mg and in the dystrophics it was 1.6±0.3 mg. Adrenal/body weight ratio

Table 1 Body and thymus weights and thymus/body weight ratios in Swiss mice and in dystrophic and phenotypically normal animals of the Bar Harbor 129 strain.

At 4 weeks of age	Swiss	Normal 129	Dystrophic 129
Number	16	8	8
Body weight (g)	21.4±1.36	17.3±0.76	9.3±1.9
Thymus weight (mg)	73.0±5.5	72.2±4.3	3.7±1.2
Thymus/body weight ratio (×10⁻²)	3.5±0.3	4.2±0.1	0.37±0.19
At 12 weeks of age	Swiss	Normal 129	Dystrophic 129
Number	16	8	8
Body weight (g)	32.4±0.9	26.2±0.82	16.5±1.1
Thymus weight (mg)	18.0±1.3	32.6±2.0	19.4±3.3
Thymus/body weight ratio (×10⁻²)	0.56±0.03	1.25±0.08	1.20±0.28

Results are shown as means ± s.e.m.

in the normals was 3.0×10^{-4} and in the dystrophics 3.1×10^{-4}. Adrenal weight is therefore a normal proportion of body weight in the dystrophic animals and it is unlikely that the thymus atrophy is caused by adrenal hyperactivity.

We then investigated the possibility that thymus growth might be affected at an earlier stage in the dystrophic animals. Three litters were killed between 15 and 20 d *post partum* before weaning. Dystrophic animals were identified by muscle histology. Thymus/body weight ratio in the normal 129s was 3.9×10^{-3} and in the dystrophics 2.3×10^{-3}. Thus the thymus was lower in weight at this stage but the severe reduction noted in the dystrophics one week after weaning was not found.

This pattern is similar to the situation in growth hormone and prolactin-deficient dwarf mice[6-8]. In these animals thymus growth is normal until weaning, apparently being maintained by a factor in the milk. On weaning rapid thymus atrophy occurs. This can be reduced or prevented by treatment with exogenous growth hormone or prolactin. Prolactin is known to be present in rodent milk[9] and it is possible that this is the necessary factor. Quite independently, de Kretser and Livett[10] in Melbourne have made similar observations on thymus growth in dystrophic mice. The interpretation of our common findings must obviously be uncertain at this stage although our own hypothesis is that both muscle and thymus are damaged by the basic genetic defect but that the clinical expression of the genetic defect in the muscles may depend on some of the consequences of thymus damage.

This work was carried out partly at the University of Newcastle upon Tyne, UK. We thank the North of England Cancer Campaign, the UK Muscular Dystrophy Association and the Quebec Medical Research Council for financial support, and Professor W. G. Bradley and the Staff of the Muscular Dystrophy Laboratories, Newcastle General Hospital, for the Bar Harbor 129 mice, and for much advice.

R. A. KARMALI
D. F. HORROBIN

Clinical Research Institute and Université de Montréal,
110 Pine Avenue West,
Montreal, Canada

1 Walker, B. E, *Tex. Rep. Biol. Med* 22, 640, (1964).
2 Matheson, D. W., *J. neurol. Sci* 81, 133, (1974).
3 Micholson, A. M., Russell, E. S., and Harman, P. J, *Proc. natn Acad. Sci. U.S.A.,* 41, 1079, (1955).
4 Stevens, L. C., Russell, E. S. and Southard, J. L., *Proc. Soc. exp Biol Med.,* 95, 161, (1957).
5 Harman, P. J., Tassoni, J. P, Curtis, R. L, and Hollinshead, M. B., in *Muscular Dystrophy in Man and Animals* (edited by Bourne, G. H., and Golarz, N), 407 (Basel, and New York, 1963).
6 Fabris, N., Pierpaoli, W., and Sorkin, E., *Clin. exp Immun.,* 9, 209, (1971).
7 Fabris, N, Pierpaoli, W, and Sorkin, E. *Clin. exp. Immun.,* 9, 227, (1971).
8 Fabris, N, Pierpaoli, W., and Sorkin, E., in *Hormones and the Immune Response* (edited by Wolstenholme, G. E W., and Knight, J.,), (Churchill, London, 1970).
9 McMurphy, J. P., and Malven, P. V , *J. Endocrinol.,* 61, 211 (1974).
10 Livett, and de Kretser, *Nature,* 263, 682-684 (1976).

Foetoneonatal oestradiol-binding protein in mouse brain cytosol is α foetoprotein

BRAIN cytosol from foetal and neonatal rodents contains high concentrations of an oestradiol-binding protein (foetoneonatal oestradiol-binding protein or FEBP) with properties distinctly different from those of adult oestrogen receptors with respect to sedimentation rate, binding specificity and affinity for oestradiol[1-3]; FEBP declines to undetectable levels by about 22 d in rat brain cytosol[1,3]. Sexual differentiation of the rodent brain, which normally occurs under the influence of testicular secretions between birth and about 10 d, can also be brought about by injection of oestradiol benzoate into neonatal female or castrated male rodents[4]. This observation has led to the idea that FEBP protects the developing brain against the effects of maternal oestrogens[5]. It has been suggested that FEBP in rats is α foetoprotein (AFP)[5], a circulating foetal protein known to bind oestradiol[6-8]. Here we present evidence which points strongly to the identity of mouse FEBP and AFP This includes a comparison of the binding properties of FEBP with those of purified mouse AFP, correlation of the postnatal disappearance of FEBP with measurements of AFP in brain cytosol and removal of oestradiol-binding activity from brain cytosol using anti-mouse AFP antibodies.

Cytosol—prepared from a homogenate of pooled brains from neonatal mice by centrifugation at 216,000g—or mouse AFP—purified by immunoadsorbents[10,11]—was incubated for 2 h at 0 °C with 5×10^{-9} M ³H-oestradiol and centrifuged on glycerol gradients (Fig. 1). The sedimentation coefficients of the peaks of ³H-radioactivity were estimated to be 4.6±0.03S for FEBP (mean ±s.e. of nine determinations involving five preparations from 1-5-d-old mice) and 4.7±0.02S for AFP (mean ±s.e. of 13 determinations on four different AFP preparations). This slight difference in the sedimentation rates of AFP and FEBP could be due to a change that occurred in the AFP molecule as a result of the purification; this was supported by the observation that an AFP preparation purified from brain cytosol gave an S value (4.7 S) comparable with that of the other purified AFP preparations. Figure 1 also shows that excess unlabelled oestradiol (10^{-6} M) decreased binding of ³H-oestradiol to both FEBP and AFP by 80–90%. Diethylstilbestrol (DES), a potent synthetic oestrogen, at 10^{-5} M, had no effect on oestradiol binding to FEBP, in agreement with results reported for rat FEBP[1,3]; 103% (mean of four determinations) of the ³H-c.p.m. found in this peak in the absence of DES was bound in its presence. DES competed slightly, however, with ³H-oestradiol for binding to purified mouse AFP (17%, mean of four determinations); ³H-DES has also been reported to bind to purified rat AFP[6]. The apparent difference in sensitivity to DES competition between FEBP and AFP was abolished when purified AFP
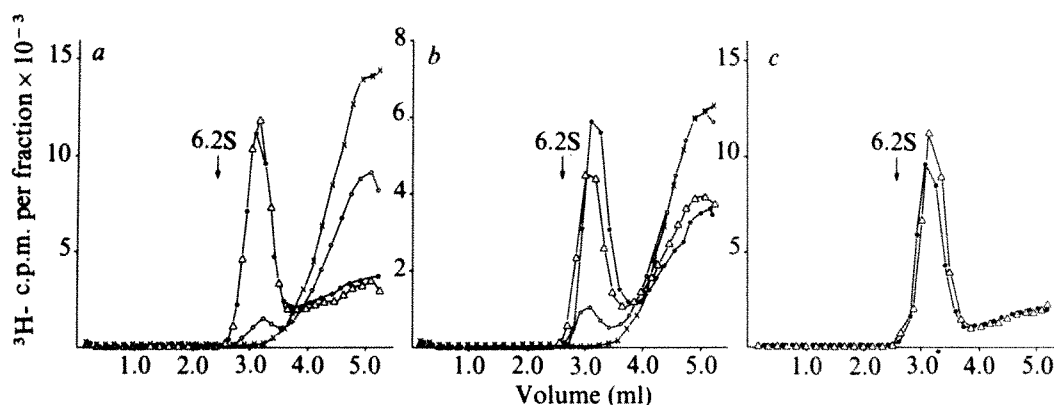
Fig. 1 Glycerol gradient centrifugation of 5-d brain cytosol and mouse AFP incubated with ^{3}H-oestradiol and ^{3}H-testosterone. Cytosol was prepared from pooled brains of 5-d-old female mice or from 3–5-month-old adult mice in TEM (0.01 M Tris, pH 6.7, at 25 °C; 0.0015 M EDTA; 0.002 M β-mercaptoethanol) by centrifugation for 1 h at 216,000g as described before[13]. Mouse AFP, purified by immunoadsorbents[10,11], was dissolved in TEM. Samples of 0.4 ml of 5-d cytosol (4.7 mg per ml of protein) (a), of AFP (18 μg ml⁻¹) (b), or of AFP (18 μg ml⁻¹) in adult cytosol (6.8 mg per ml of protein) (c) were incubated with 5 × 10⁻⁹ M [3]H-17β-6,7-oestradiol (47.9 Ci mmol⁻¹) in the absence (●) or presence of unlabelled oestradiol at 10⁻⁶ M (○) or DES at 10⁻⁶ M (△), and with 5 × 10⁻⁹ M [3]H-1α,2α-testosterone (59 Ci mmol⁻¹) (×) for 2 h at 0 °C. Samples were layered on 5–35% linear glycerol gradients in TEM and centrifuged in the SW 50.1 rotor at 216,000g for 16 h at 2 °C. The scale on the left of (b) refers to the gradients containing ^{3}H-oestradiol, while that on the right refers to the gradient containing ^{3}H-testosterone. The arrow indicates the position of bacterial alkaline phosphatase added as an internal sedimentation marker (6.2S, see ref. 13). The direction of sedimentation was from right to left.

was incubated with ^{3}H-oestradiol in the presence of brain cytosol from adult mice (which contains no detectable FEBP). DES no longer displaced any radioactivity from the AFP peak (Fig. 1); in two such experiments, 104% and 109% of the ^{3}H-c.p.m. were recovered in the AFP peak in the presence of DES. There was no indication of binding of ^{3}H-testosterone to either FEBP or mouse AFP (Fig. 1).

Binding parameters for the reaction of brain cytosol or mouse AFP with ^{3}H-oestradiol were obtained by Scatchard[12] analysis after equilibrium dialysis. This method was chosen since we[13] and others[14] have observed that there is a rapid dissociation of the FEBP-oestradiol complex in non-equilibrium conditions. The mean dissociation constants (k_D) from two experiments were 2.3 × 10⁻⁸ M (brain FEBP) and 2.4 × 10⁻⁸ M (mouse AFP); these are essentially the same as published values for rat brain FEBP[1] and purified rat AFP[9].

AFP was demonstrable in brain cytosol from immature mice by immunodiffusion or radioimmunoassay[10]. Cytosol from 1-, 4–5-, 9–10- and 20–22-d-old mice gave AFP concentrations (mean ±s.e.) of 7.0±0.8 μg (seven determinations), 1.5±0.2 μg (three determinations), 0.24±0.02 μg

Fig. 2 Scatchard plots representing binding of ^{3}H-oestradiol by 4-d brain cytosol (a) and mouse AFP (b). Samples of 0.35 ml of brain cytosol from 4-d-old male mice (11 mg per ml of protein) or of AFP (17 μg ml⁻¹) were dialysed against 10 ml TEM containing various concentrations of ^{3}H-oestradiol (OII) (2 × 10⁻⁹–1 × 10⁻⁷ M) for 24 h at 4 °C. Triplicate samples of the inner and outer solutions were removed for counting. Bound radioactivity was determined from the difference between the radioactivity inside the dialysis bags and that outside.

(three determinations), and <2 ng (undetectable level) per mg of protein, respectively. There were 0.7–0.8 mol of oestradiol bound per mol of mouse AFP (molecular weight 74,000 (ref. 10)), and the binding capacity of brain cytosol from 4-d-old mice was about 25 pmol per mg of protein (Fig. 2). Assuming the binding moiety in brain cytosol to be AFP with one binding site per molecule, 4-d brain cytosol should contain 1.8–1.9 μg of

Fig. 3 Glycerol gradient analysis of brain cytosol from mice of various ages incubated with ^{3}H-oestradiol. Cytosol from two or three female mice aged 1 (●), 4 (×), 10 (○) and 21 d(△) was incubated with 2 × 10⁻⁸ M ^{3}H-oestradiol. Samples of 0.4 ml containing 1.6, 2.6, 3.2, and 3.7 mg of protein, respectively, were centrifuged on glycerol gradients as in Fig. 1.

AFP per mg of protein. This is in close agreement with the observed AFP concentration (1.5 μg per mg of protein) at this age. The time course of disappearance of 4.6S oestradiol-binding activity in brain cytosol (Fig. 3) parallels the decrease in AFP levels. At 1 and 5 d substantially all the ^{3}H-radioactivity was found in the 4.6S peak (at 5×10[-9] ^{3}H-oestradiol). At 10 d there was an appreciable decrease in the proportion of bound ^{3}H-oestradiol, with a concomitant increase in free steroid at the top of the gradient, and by 21 d no binding was detectable.
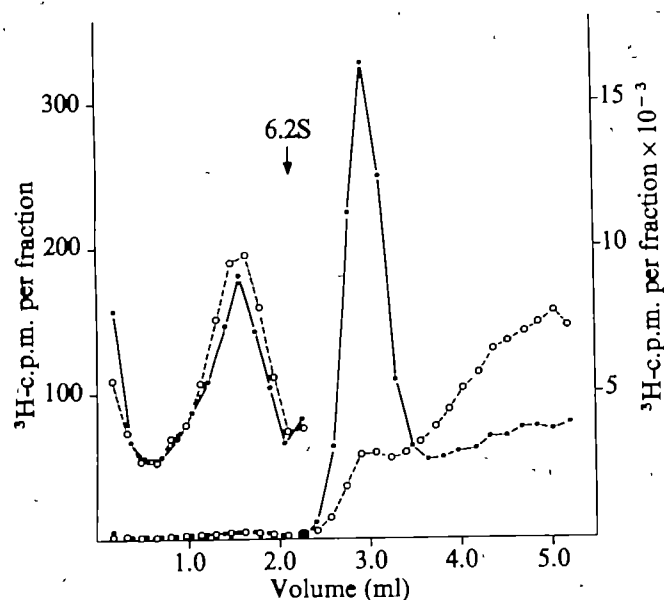
Fig. 4 Glycerol gradient analysis of brain cytosol from 5-d-old male mice incubated with ^{3}H-oestradiol after treatment with anti-mouse AFP or an unrelated control adsorbent. Cytosol was treated with Sepharose-coupled anti-mouse AFP antibodies (○) or with control adsorbent (●) for 3 h at 0 °C and incubated with 5×10[-9] M ^{3}H-oestradiol. Samples of 0.5 ml containing 2.5 and 2.2 mg of protein, respectively, were centrifuged on glycerol gradients as in Fig. 1. The scale on the left refers to the expanded plot of the bottom part of the gradients showing the peaks of 8S oestradiol receptor.

Treatment of 5-d brain cytosol with anti-mouse AFP antibodies insolubilised on to Sepharose removed almost completely the 4.6S oestradiol-binding activity, while this binding was not affected in a sample treated with an unrelated control Sepharose adsorbent (Fig. 4). The binding data were in agreement with measurements of AFP in these samples: these were 0.12 and 1.4 μg of AFP per mg of protein, respectively. In contrast, there was no change in the oestradiol-binding capacity of the 8S oestradiol receptor[13] after exposure of cytosol to anti-mouse AFP. A second experiment involving 1-d brain cytosol gave essentially the same results.

Thus FEBP in neonatal mouse brain cytosol has properties similar to those of purified mouse AFP with respect to affinity for 17β-oestradiol, specificity of the binding reaction and sedimentation rate in glycerol gradients. In addition, AFP can be detected in cytosol preparations by immuno-diffusion or radioimmunoassay, and AFP and 4.6S oestradiol-binding activity show a similar time course of disappearance in the developing brain and coincident removal from brain cytosol by anti-mouse AFP antibodies. By all these criteria AFP and FEBP are substantially identical.

The biological role of AFP is unclear, but its high capacity for oestradiol (and oestrone) has led to the idea that it may be involved in the protection of foetal tissues against circulating maternal oestrogens[1,4]. This may be particularly

important in the brain because the FEBP-AFP concentra-tion in brain cytosol seems to be higher than can be accounted for by contamination from the blood[1,4]. The oestrogen-binding properties of AFP, however, may be confined to rodents and this may not be the only function of this protein. This is suggested by the lack of significant oestrogen binding, demonstrable by equilibrium dialysis or gradient centrifugation, by human AFP-containing serum or amniotic fluid[15,16] or by purified human and bovine AFPs (our unpublished results). Also AFP shows sequence homo-logy with serum albumin[17], and oestrogen binding may be only a specialised expression of general albumin-like binding properties of the AFP molecule.

BARBARA ATTARDI

Department of Biology,

ERKKI RUOSLAHTI

Department of Immunology,
City of Hope National Medical Center,
Duarte, California 91010

1 Plapinger, L., McEwen, B. S., and Clemens, L. E., *Endocrinology*, 93, 1129–1139 (1973).
2 Kato, J., Atsumi, Y., and Inaba, M., *Endocrinology*, 94, 309–317 (1974).
3 Fox, T. O., *Nature*, 258, 441–444 (1975).
4 Goruki, R. A., *J. Reprod. Fert.*, Suppl., 1, 67–88 (1966).
5 McEwen, B. S., Plapinger, L., Chaptal, C., Gerlach, J., and Wallach, G., *Brain Res.*, 96, 400–406 (1975).
6 Plapinger, L., and McEwen, B., *Steroids*, 26, 255–265 (1975).
7 Savu, L., et al., *FEBS Lett.*, 22, 113–116 (1972).
8 Uriel, J., de Nechaud, B., and Dupiers, M., *Biochem biophys Res Commun.* 46, 1175–1180 (1972).
9 Aussel, C., Uriel, J., and Mercier-Bodard, C., *Biochimie*, 55, 1431–1437 (1973).
10 Pihko, H., and Ruoslahti, E., *Int. J. Cancer*, 12, 354–360 (1973).
11 Ruoslahti, E., Pihko, H., and Seppala, M., *Transplant Rev.*, 20, 38–60 (1974).
12 Scatchard, G., *Ann. N.Y. Acad. Sci*, 51, 660–672 (1949).
13 Attardi, B., and Ohno, S., *Endocrinology* (in the press).
14 Barley, J., Ginsburg, M., Greenstein, B. D., MacLusky, N. J., and Thomas, P. J., *Nature*, 252, 259–260 (1974).
15 Nunez, E., Vallotl, G., Bonassaya, C., and Jayle, M. F, *Biochem biophys. Res. Commun.*, 57, 126–133 (1974).
16 Swartz, S. K., Soloff, M. S., and Surtano, R. R., in *Colloque on Alphafeto-protein* (edit. by Masseyeff, R.), 97–108 (INSERM, Paris, 1974).
17 Ruoslahti, E., and Terry, W. D., *Nature*, 260, 804–805 (1976).

Inhibition of platelet aggregation by a myeloma protein with anti-phosphocholine specificity

PHOSPHOCHOLINE-containing phospholipids, particularly phosphatidylcholine (lecithin), lysophosphatidylcholine and sphingomyelin, have integral roles in the structure and function of mammalian cell membranes[1]. Enzymatic diges-tion of platelet choline phosphatides, particularly of phos-phatidylcholine which is the most abundant of the platelet phospholipids[2,3], has been shown to induce platelet release reactions[4]. We have demonstrated that C-reactive protein (CRP), an acute phase reactant with binding specificities for phosphocholine, choline phosphatides and polycations generally[5,6], can inhibit multiple platelet reactivities[7,10]. These considerations have raised the possibility that mole-cules generally with binding specificities for choline phos-phatides can influence platelet responsiveness. We report here an investigation of the ability of the T-15 anti-phospho-choline mouse IgA myeloma protein to inhibit the aggre-gating response of human platelets.

Diluents and human platelets were prepared and platelet viability and aggregation were assayed as described before[8]. The IgA-producing plasma-cell tumours TEPC-15 and MOPC-315 were provided by Dr Donald A. Rowley, Univer-sity of Chicago, and were maintained in BALB/cJ mice (Jackson Laboratories). The anti-phosphocholine myeloma protein T-15 was isolated by affinity chromatography from pools of ascites fluid obtained from tumour-bearing mice essentially by the method of Chesebro and Metzger[11], except
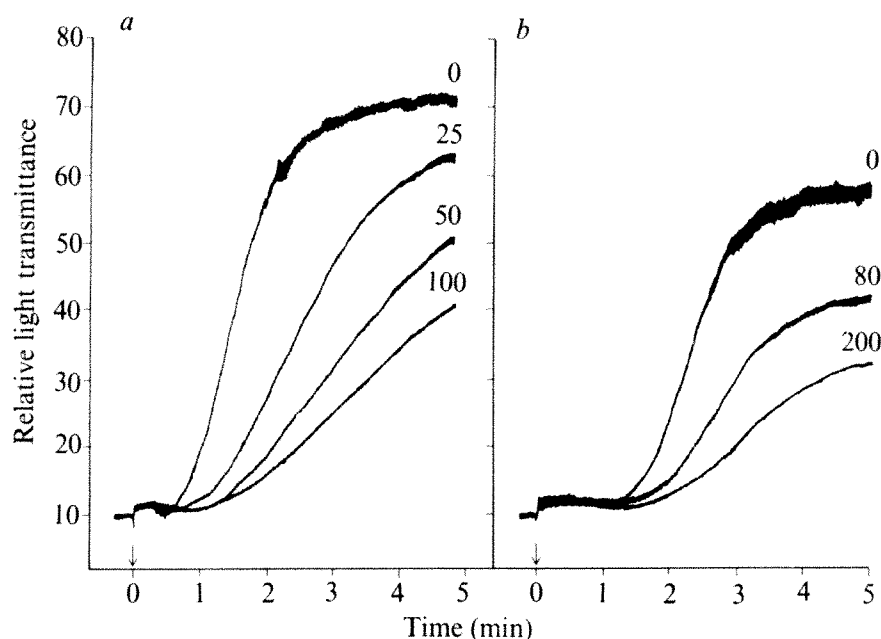
Fig. 1 The effects of increasing amounts of T-15 on platelet aggregation induced by thrombin (0.2 units ml^{-1}) (a) and AHGG (50 µg ml^{-1}) (b); the concentrations of T-15 added to the reaction mixtures are expressed in µg ml^{-1}.

that C-polysaccharide-Agarose beads[12] were used in place of phosphocholine-Agarose. The protein was purified further by filtration on BioGel A-0.5 m, and the material reactive with rabbit anti-mouse IgA antiserum corresponding to the 11S dimer was used. Similarly, the anti-dinitrophenol myeloma protein M-315 was isolated from pools of serum and ascites fluid obtained from mice bearing MOPC-315 by affinity chromatography on trinitrophenylaminoethyl Agarose beads (BioGel, A-15 m); it was eluted with 0.1 M dinitrophenol, and the dinitrophenol was removed on Dowex 1 (ref. 13).

The T-15 mouse IgA myeloma protein reactive with phosphocholine was found to inhibit the ability of washed human platelets to aggregate in response to thrombin and aggregated human γ globulin. This inhibition was characterised by a reduced rate (slope) and extent (amplitude) of aggregation, and a decreased size of the platelet aggregates formed (magnitude of the vertical oscillations); concentrations as low as 25–80 µg ml^{-1} were sufficient to bring about these effects (Fig. 1).

By contrast, the M-315 mouse IgA myeloma protein did not inhibit aggregation at concentrations as high as 425 µg ml^{-1}, indicating that the inhibition by T-15 involved its antigen-combining site. In support of this interpretation, inhibition by T-15 was reduced substantially when phosphocholine was included in the reaction mixtures (Table 1).

Table 1 Effect of phosphocholine on T-15-inhibited aggregation of platelets stimulated by thrombin and AHGG

Inhibitor	% Inhibition ± s.d.*	
	Thrombin	AHGG
T-15	56 ± 10.0	54 ± 7.1
T-15 + PC	15 ± 7.1	26 ± 1.4
PC	1 ± 0.5	2 ± 1.4

Platelets (3 × 10⁹) were equilibrated with T-15 for 1 min at 37 °C and challenged with thrombin (0.2 units ml^{-1}) or AHGG (50 µg ml^{-1}); aggregation was recorded for 5 min. The concentration of T-15 chosen was sufficient to inhibit the rate of aggregation by approximately 50% and represents 60 (thrombin) and 80 (AHGG) µg ml^{-1}; phosphocholine (PC) was at a final concentration of 10⁻³ M.
*The data are presented as the percentage inhibition in the rate of platelet aggregation ± 1 standard deviation (s.d.) as compared to the aggregation response obtained in the absence of inhibitors.

The possibility that T-15 acts by reducing platelet viability was discounted because concentrations far in excess of those used in this study (>950 µg ml^{-1}) were not cytotoxic as determined by incorporation of ^{14}C-5-hydroxytryptamine (^{14}C-5-HT) and release of lactic dehydrogenase.

Thus, the T-15 myeloma protein with anti-phosphocholine specificity inhibits the reactivity of human platelets with characteristics similar in all respects investigated to those previously observed when CRP was added to these mixtures. It had not been clear whether CRP reacts with phosphocholine lipids, polycations or other of its binding specificities to cause its inhibitory effect. Our data suggest that a phosphocholine moiety on the platelet surface is involved in this inhibition, and that phosphocholine-binding molecules generally can modulate platelet reactivities; CRP would seem to be biologically the most important of the latter. The anti-phosphocholine myeloma protein may well aid in the delineation of the surface receptors and pathways involved in the CRP-mediated control not only of platelet reactivities, but also of responses of other cells reactive with CRP, such as subpopulations of T cells[14-16]. The nature of these receptors, and the significance of the association of molecules with which they interact with inflammation and tissue destruction, seem to be critical to a more complete understanding of body defence mechanisms.

Barry A. Fiedel
Alexander P. Osmand
Henry Gewurz

Department of Immunology,
Rush Medical College and Department of
Microbiology,
University of Illinois at the Medical Center,
Chicago, Illinois 60612

1 Finean, J. B., in *Form and Function of Phospholipids*, 171 (Elsevier, New York, 1973).
2 de Gier, J., de Haas, G. H., and Van Deenen, L. L. M., *Biochem. J.*, 81, 33P (1961).
3 Marcus, A. M., Ullman, H. L., and Safier, L. B., *J. Lipid Res.*, 10, 108 (1969).

4 Schick, P K , and Yu, B. P ,*J. clin Invest* , 54, 1032 (1974).
5 Gotschlich, E. C., and Edelman, G. M , *Proc natn Acad Sci U.S A* , 57, 706 (1967).
6 Volanakis, J. E , and Kaplan, M H.,*Proc. Soc exp Biol Med* ,236, 612 (1971)
7 Kaplan, M H., and Volanakis, J. E.,*J. Immun* ,112, 2135 (1975)
8 Siegel, J , Osmand, A. P., Wilson, M. F , and Gewurz, H ,*J. exp Med* , 142, 709 (1975)
9 Fiedel, B. A., and Gewurz, H.,*J. Immun.*, 116, 1289 (1976).
10 Fiedel, B A., and Gewurz, H ,*J Immun* (in the press).
11 Chesebro, B , and Metzger, H., *Biochemistry*, 11, 766 (1972)
12 Osmand, A. P , Mortenson, R F., Siegel, J., and Gewurz, H ,*J exp Med* , 142, 1065 (1975)
13 Eisen, H N , Gray, W., Little, J. R , and Simms, E. S , in *Methods in Immunology and Immunochemistry*, 1, 351 (Academic, New York, 1967)
14 Mortenson, R F., Osmand, A P., and Gewurz, H .,*J. exp Med* , 141, 821 (1975).
15 Mortensen, R. F., and Gewurz, H ,*J Immun.*, 116, 1244 (1976)
16 Croft, S. M., Mortenson, R. F , and Gewurz, H , *Science*, 193, 685 (1976)

Antiviral antibodies inhibit the lysis of tumour cells by anti-H-2 sera

RECENT experiments have indicated that the cytolysis of virus-infected, chemically modified, or neoplastic cells by murine T lymphocytes is restricted to target cells carrying one or more of the H-2 antigens possessed by the cells against which the cytotoxic T cells were generated[1-10]. Two general hypotheses have been proposed to explain the participation of H-2 antigens in T-cell-mediated cytolysis. According to one hypothesis, T lymphocytes possess two different types of receptors, one which recognises the abnormal antigens of the target cell, and a second which interacts with the H-2 antigens of that cell. In the alternative hypothesis, the T cell has only one specific receptor that is directed against a complex antigenic determinant made of of H-2 antigens and viral antigens, for example, tumour antigens induced by viral infection. Neither of these hypotheses has yet been satisfactorily proven. In support of the one-receptor hypothesis, we have found that anti-H-2 and antiviral sera show copatching and cocapping of H-2 and viral antigens on the surface of EL4 leukaemia cells[a], suggesting that viral and H-2 antigenic determinants can be in close physical association on the cell surface. Here we provide new evidence for the existence of a close association between H-2 antigens and viral antigens on murine leukaemia cells.

We have used the method of antibody-induced resistance to complement-dependent lysis ("lysostrip" technique[11]) to show that the binding of antiviral antibody to the cell surface resulted in resistance to cell lysis induced by subsequent treatment with anti-H-2 serum and complement. Bernoco *et al*.[11] and Hauptfeld *et al*.[12] have shown that specific resistance to lysis by an alloantiserum can be effected by first coating cells with the alloantiserum and then incubating the cells with an anti-Ig serum directed against the first antibody.

Treatment of EL4 cells (H-2ᵇ) with a rabbit antiserum to Rauscher leukaemia virus (RLV), or with a goat antiserum against the purified glycoprotein gp70 of murine leukaemia virus (Scripps), and then with an appropriate anti-Ig reagent reduced the subsequent complement-mediated lysis of the cell by anti-H-2ᵇ antiserum (Table 1). In contrast, both of the antiviral antisera had only slight and inconsistent effects in similar experiments using normal lymphocytes from the parental mouse strain C57BL/6 (Table 1). The corresponding control sera (normal rabbit serum taken before immunisation with RLV or pooled normal goat serum) had no effect either on EL4 cells or spleen cells. Before use, all of the heterologous sera were absorbed at least twice with equal volumes of normal spleen cells.

The efficiency and specificity of the "lysostrip" technique was demonstrated on both normal spleen cells and tumour cells. Pretreatment with antisera to the H-2Dᵇ antigens caused no reduction or very little reduction in the

susceptibility to lysis by an anti-H-2Kᵇ antisera (Table 1). Similarly, precoating the cells with antiserum to the H-2Kᵇ antigens caused no reduction in susceptibility to lysis by anti-H-2Dᵇ serum (Table 1) Precoating the cells with an anti-H-2Kᵇ antiserum, however, did lead to an increased lysis in the cytotoxicity assay performed with an anti-H-2Dᵇ antiserum. This increase was probably due to anti-H-2Kᵇ antibodies that had bound to the cells and were present during the cytotoxic assay. It was not observed if the precoated cells were treated with goat antibody to mouse IgG, which also suggests that it was due to anti-H-2Kᵇ antibodies that were bound in the first step of the experiment.

These results confirm that resistance to lysis can be induced only by redistribution of the H-2 antigen specifically tested in the cytotoxicity assay, as reported by Hauptfeld *et al.*[11]. The observation that the pretreatment of EL4 cells with antiviral sera induced a partial resistance to cytotoxicity mediated by alloantisera against both the K and the D end of the major H-2 complex is therefore not likely to be a peculiarity of the method that we used.

The identity of the relevant viral antigens carried by the EL4 cells is unknown Sodium dodecyl sulphate (SDS)-polyacrylamide gel electrophoresis of immune precipitates
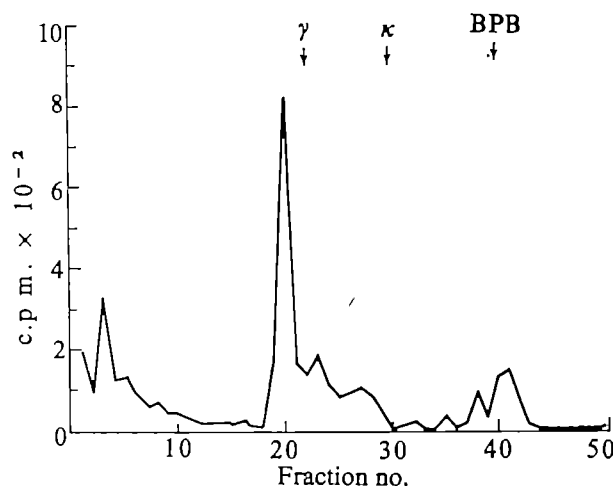
Fig. 1 SDS–polyacrylamide gel electrophoresis of an immuno-precipitate of an extract of ¹²⁵I-surface-labelled EL4 cells. EL4 cells were iodinated with ¹²⁵I by the lactoperoxidase technique[21] and extracted with 0.5% NP40 (ref. 19). Immuno-precipitates were obtained using BALB/c absorbed RLV and unabsorbed GARIgG. SDS gel electrophoresis was performed in 8 5% polyacrylamide gels after solubilisation and reduction of the precipitates by boiling for 5 min in sample buffer containing 4% SDS and 5% mercaptoethanol[19]. MOPC 21 myeloma immunoglobulin labelled with ¹³¹I was used as marker. γ, Heavy chain (molecular weight, 50,000);κ, light chain (molecular weight, 23,000); BPB, bromophenol blue marker.

made from ¹²⁵I-labelled EL4 cells using RLV, however, showed a major component corresponding in mobility to the viral glycoprotein gp70 (Fig. 1). Similar results were obtained with goat antisera against the gp70 glycoproteins of Rauscher leukaemia virus[13] or murine leukaemia virus (Scripps). All three antiviral antisera also stained the EL4 cells brightly by immunofluorescence. Viral antigens were not detected on C57BL/6 spleen cells with these antisera by either immunofluorescence or immune precipitation.

Of the artefacts that might arise in these experiments, the most obvious would be ascribable to antiviral antibodies in the anti-H-2 antisera[13-15] or anti-H-2 antibodies in the antiviral antisera. Several kinds of control experiments were performed to explore these possibilities. Most

Table 1 Effect of antiviral sera on cytolysis by anti-H-2 sera

Antisera used in:—		EL4								C57BL/6		
Step 1	Step II	K^b*	K^b	K^b	K^b	K^b	K^b†	D^b	D^b	K^b	K^b	D^b
None	None	33	25	33	45	35	61	23	22	34	41	23
RARLV	None	35			45			30	28	34	41	30
RARLV	GARIgG	21			34			19	15	33	39	35
NRS	None	31						29	29	28		37
NRS	GARIgG	33						30	32	28		37
None	GARIgG	39						28	26	29	40	37
GAgp70	None		22	29	40	36				28		
GAgp70	RAGIgG		12	16	16	26				33		
NGS	None		23	33						24		
NGS	RAGIgG		27	31						29		
None	RAGIgG			33		34				23		
Anti-H-2D^b	None						71	33	32	34		20
Anti-H-2D^b	GAMIgG						65	16	13	30		3
Anti-H-2K^b	None	39	27	34	49	35	61	42	39	37	40	37
Anti-H-2K^b	GAMIgG	11	8	12	9	15	39	33	30	8	−1	22
None	GAMIgG	26				43		29	23	31	39	23

RLV, rabbit antiserum against Rauscher leukaemia virus; GARIgG, goat antiserum against- rabbit IgG; NRS, normal rabbit serum; GAgp70, goat antiserum against gp70; RAGIgG, rabbit antiserum against goat IgG; NGS, normal goat serum; GAMIgG, goat antiserum against mouse IgG. The procedure was performed in three steps. Step I. 5×10^5 ^{51}Cr-labelled cells were coated with 0.02 ml antiviral or anti-H-2 antibody in 0.2 ml Minimal Essential Medium with 5% heat-inactivated foetal bovine serum (MEM/FBS) for 30 min on ice. Step II. After washing twice, cells were incubated for 45 min at 37 °C in 0.2 ml MEM/FBS with 0.02 ml of a second antiserum directed against the first antibody. Cells were washed once and this step was repeated. After completion of this second cycle cells were washed once and resuspended in 0.22 ml MEM/FBS. Step III. Cytotoxicity assay was performed using 2×0.05 ml of this cell suspension as described earlier[19] using antisera directed against the H-2K^b or H-2D^b antigens as indicated (*). Total c.p.m. were determined in 2×0.05 ml of the same cell suspension. Results are given in percentage specific lysis calculated according to the following formula:

$$\frac{\text{Experimental c.p.m.} - \text{background c.p.m.}}{\text{Total releasable c.p.m.} - \text{background c.p.m.}} \times 100$$

Background lysis was determined separately and varied between 14 and 30%. Each antiserum used was also tested individually to exclude nonspecific lysis due to incomplete absorption. EL4 cells (Salk Institute) were grown in ascitic form in C57BL/6 mice (Jackson Laboratory). Splenic lymphocytes were prepared by the Ficoll-Hypaque technique[20]. Cells were labelled with 100 μCi of ^{51}Cr per 10^7 cells. Antisera used in step I were anti-H-2K^b (H-2.33) and anti-H-2D^b (H-2.2) prepared as described earlier[19] and were used undiluted. RLV was produced by repeated injections of SDS-solubilised Rauscher leukaemia virus (Litton Bionetics) adding complete Freund's and incomplete Freund's adjuvant, respectively, in the first two injections. RLV was absorbed twice with equal volumes of packed C57BL/6 spleen cells or with BALB/c spleen cells (H-2^d) when used on C57BL/6 spleen lymphocytes. NRS was obtained from the rabbit producing the RLV before the first immunisation and was absorbed by the same procedure. GAgp70 raised against gp70 isolated from murine leukaemia virus (Scripps), and pooled NGS were absorbed in the same way as the rabbit sera. Antisera used in step II were GAMIgG (Miles-Yeda, Rehovot) was absorbed with BALB/c thymus cells. GARIgG prepared by repeated injections of rabbit IgG was absorbed with BALB/c spleen and thymus cells, and RAGIgG (Cappel Laboratories) was absorbed with BALB/c spleen cells. In the cytotoxic assay the antiserum against H-2K^b was used at a dilution of 1:18 (except K^b†, dilution: 1:12) and the antiserum against H-2D^b was used at 1:8. Guinea pig complement (Grand Island Biologic. Co., Grand Island) was absorbed with mouse spleen cells and diluted 1:4. Each vertical column represents a single experiment. Results indicating that pretreatment has reduced the susceptibility to complement-mediated lysis by anti-H-2 sera are underlined. When anti-H-2D^b serum was used in the cytotoxicity assay, there was increased specific lysis of cells that had been pretreated with RLV or anti-H-2D^b serum alone. For this reason, the lysostrip effect is probably best estimated by comparing the specific lysis of cells pretreated with the first antibody and then anti-Ig antibodies, with the specific lysis of cells pretreated only with the first antibody or only with anti-Ig antibodies.

mouse alloantisera contain antibodies directed against murine leukaemia antigens[14,15] and particularly against the envelope protein gp70 (refs 13-15). The cytotoxic activity of certain alloantisera against tumour cells may therefore consist of two components, one due to antibodies against alloantigens and the other to antibodies against viral antigens. Accordingly, the anti-H-2^b sera used in this study were absorbed three times with equal volumes of packed thymocytes of young BALB/c mice (H-2^d) or with NZB-thymoma cells that were actively producing murine leukaemia virus (Scripps) in large amounts (SCRF 60A, H-2^d). After absorption, the titres of the anti-H-2K^b and anti-H-2D^b sera were compared using EL4 cells as targets. There was no difference in the cytotoxic titres after absorption with either normal thymocytes or virus-producing tumour cells.

This result indicated that our cytotoxicity assay did not detect anti-viral antibodies in the anti-H-2 antisera. It is therefore unlikely that the decreased cytotoxic effect of the anti-H-2 antisera after redistribution of viral antigens merely reflects the absence of cytolysis that might have been due to the presence of antiviral antibodies in the anti-H-2 antisera. It should be noted that small amounts of anti-viral antibodies were present in the anti-H-2D^b antiserum, however, inasmuch as a gp70 component was

observed in the SDS-gel electrophoretic analysis of immune precipitates made from ^{125}I-labelled EL4 cells or C57BL/6 spleen cells[13]. On the other hand, immune precipitates using the anti-H-2K^b antiserum on extracts of EL4 or C57BL/6 cells did not contain detectable amounts of viral antigens[13].

The specificity of one of the anti-viral antisera used (RLV) was demonstrated by comparing the "lysostrip" effect of RLV on EL4 cells after absorption with either C57BL/6 spleen cells or with the cell line SCRF 60A (H-2^d) that produces murine leukaemia virus (Scripps). Pretreatment with 4 or 0.8 μl of the RLV that had been absorbed with SCRF 60A (H-2^d) cells did not reduce subsequent cytolysis by H-2K^b antiserum in contrast to pretreatment with similar amounts of the RLV that had been absorbed with C57BL/6 spleen cells (Table 2). The difference between the absorption with virus-producing cells and with normal spleen cells suggests that the effect of RLV in the lysostrip assay is due to antibodies directed against viral antigens rather than to antibodies against H-2 antigens. Similar experiments showed that absorption of the H-2K^b antiserum with C57BL/6 spleen cells completely abolished its ability to induce resistance against anti-H-2K^b mediated cytotoxicity.

These observations confirm and extend our previous

observations on the cocapping of H–2 and viral antigens[8] using the same cell type (EL4) and similar antisera. Both sets of data are explicable in terms of a close physical association or inter-action of viral and H–2 determinants on the cell surface. Because the exact mechanism of the lysostrip phenomenon is not understood, however, the present results are not conclusive evidence for the close molecular association of viral and H–2 determinants in all conditions. Obviously, more direct approaches are required to demonstrate an association at the molecular level.

Although the evidence that viral and H–2 determinants may be associated on the cell surface is only indirect, it is difficult to account for by any other means. It is therefore pertinent to consider the possible mechanism of this molecular association. One possibility is that both viral and H–2 determinants are on one molecule[16]. There are no biochemical data to favour this model, but it cannot be completely excluded at present. We have observed, however, that target cells coated with inactivated Sendai virus are killed by specific cytotoxic T cells within the first 30 min after the cells are coated with the virus[2]. These experiments suggest that any synthetic modification of the H–2 antigens that might occur must be induced very rapidly, a finding that further weakens the hypothesis.

A second possibility, which we favour[9], is that viral and H–2 antigens are in close physical association on the cell membrane. The molecular interaction between the H–2 and viral antigens could be related to a specialised function of either of the molecules. For example, H–2 antigens

Table 2 Effect of absorption of RLV with a cell line producing murine leukaemia virus (Scripps)

RLV (μl)	% Specific lysis of targets pretreated with RLV absorbed with	
	C57BL/6 spleen	SCRF 60 A
None	35	35
20	18	19
4	25	35
0.8	27	35
0.4	38	36

RLV was absorbed twice with equal volumes of packed SCRF 60A cells (grown in RPMI 1640 medium with 5% FBS) or with C57BL/6 spleen cells (H–2[b]). The lysostrip technique was performed on EL4 cells as described in the legend of Table 1 using the indicated volumes of the absorbed RLV sera in the first step, and GARIgG in the second step. Sensitivity to cytotoxicity was assayed using anti-H–2K[b] (1:18). Specific lysis was calculated as described in Table 1. The specific lysis of control cells pretreated only with RLV and not with GARIgG was 43% for the RLV absorbed with C57BL/6 cells, and 40% for the RLV absorbed with SCRF 60A cells.

could have a broad binding specificity for a wide range of viral antigens or could serve as receptors for viruses. Alternatively, H–2 antigens may have a tendency to bind weakly to many surface proteins. This might account for the finding that T-cell-mediated cytotoxicity against minor histo-compatibility antigens also seems to involve the H–2 antigens[17,18]. Although the direct demonstration of hybrid antigens does not in itself exclude the hypothesis that the T cell has two separate recognition sites, one for H–2 and the other for viral antigens, a variety of experiments favours the idea that the T cell has a single kind of receptor for recognition of hybrid antigens.

murine leukaemia virus protein gp70, and for NZB-thymoma cells producing murine leukaemia virus (Scripps).

ROLAND HENNING
JOHN W. SCHRADER
GERALD M. EDELMAN

The Rockefeller University,
New York, New York 10021

[1] Doherty, P. C., Blanden, R. V., and Zinkernagel, R. M., *Transplant. Rev.*, 29, 89–124 (1976).
[2] Schrader, J. W., Henning, R., Milner, R. J., and Edelman, G. M., in *XLI Cold Spring Harb. Symp. quant. Biol.* (in the press).
[3] Shearer, G. M., Rehn, T. G., and Garbarino, C. A., *J. exp. Med.*, 141, 1348–1346 (1975).
[4] Gomard, E., Duprez, V., Henin, Y., and Levy, J. P., *Nature*, 260, 707–709 (1976).
[5] Pfizenmaier, K., Starzinksi-Powitz, A., Rodt, H., Röllinghoff, M., and Wagner, H., *J. exp. Med.*, 143, 999–1004 (1976).
[6] Zinkernagel, R. M., *Nature*, 261, 139–141 (1976).
[7] v. Boehmer, H., and Haas, W., *Nature*, 261, 141–142 (1976).
[8] Germain, R. N., Dorf, M. E., and Benecerref, B., *J. exp. Med.*, 142, 1023–1028 (1975).
[9] Schrader, J. W., Cunningham, B. A., and Edelman, G. M., *Proc. natn. Acad. Sci. U.S.A.*, 72, 5066–5070 (1975).
[10] Schrader, J. W., and Edelman, G. M., *J. exp. Med.*, 143, 601–614 (1976).
[11] Bernoco, D., Cullen, S., Scudeller, G., Trinchieri, G., and Ceppellini, R., in *Histocompatibility Testing* (edit. by Dausset, J., and Colombani, J.), 527–537 (Munksgaard, Copenhagen, 1973).
[12] Hauptfeld, V., Hauptfeld, M., and Klein, J., *J. exp. Med.*, 141, 1047–1056 (1975).
[13] Milner, R. J., Henning, R., and Edelman, G. M., *Eur. J. Immun.* (in the press).
[14] Klein, P. A., *J. Immun.*, 115, 1254–1260 (1975).
[15] Nowinski, R. C., and Klein, P. A., *J. Immun.*, 115, 1261–1268 (1975).
[16] Blanden, R. V., Hapel, A. J., and Jackson, D. C., *Immunochemistry*, 13, 179–191 (1976).
[17] Bevan, M. J., *Nature*, 256, 419–421 (1975).
[18] Gordon, R. D., Simpson, E., and Samelson, L. E., *J. exp. Med.*, 142, 1108–1120 (1975).
[19] Henning, R., Milner, R. J., Reske, K., Cunningham, B. A., and Edelman, G. M., *Proc. natn. Acad. Sci. U.S.A.*, 73, 188–122 (1976).
[20] Boyum, A., *Scand. J. clin. lab. Invest.*, 21, 1–109 (1968).
[21] Marchalonis, J. J., Cone, R. E., and Santer, V., *Biochem. J.*, 124, 921–927 (1971).

GABA-mediated control of rat neostriatal tyrosine hydroxylase revealed by intranigral muscimol

BLOCKADE of central dopamine (DA) receptors by neuroleptic drugs is associated with an increase in the firing rate of nigro-neostriatal DA neurones[1] and an enhanced turnover rate of DA in neostriatum (caudate–putamen)[2,3].
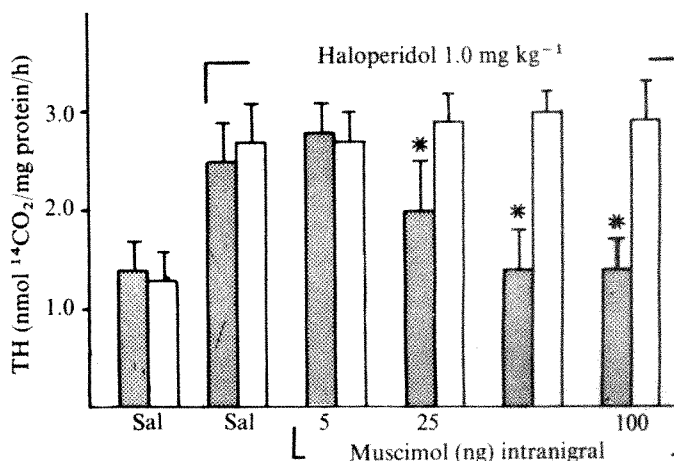
Fig. 1 Tyrosine hydroxylase (TH) activity from the right (open bars) and left (darkened bars) caudate–putamen of rats treated with saline (Sal) or muscimol directly into the left substantia nigra by way of chronically implanted stainless-steel cannulae. Muscimol was dissolved in saline and administered in a volume of 0.3 μl, over a 3-min period. Ten minutes after intranigral injection, haloperidol (1.0 mg kg[−1] dissolved in 0.05% acetic acid, intraperitoneally) or its vehicle alone (first pair of bars) was administered intraperitoneally. Animals were decapitated 40 min after intranigral injection and the tissue to be assayed was rapidly dissected out and frozen on dry ice. Assay for TH was done according to the method of Zivkovic et al.[3]. Values shown were obtained using 0.3 mM DMPH$_4$ cofactor and 0.1 mM tyrosine. Each value represents the mean of eight rats (bar indicates s.e.). *$P < 0.05$ when compared with the contralateral (control) side.

At the molecular level, the neuroleptic-induced increase of caudate–putamen (CP) DA metabolism is maintained by a decrease in the K_m of tyrosine hydroxylase (TH), for pteridine cofactors[4-6]. This increase in affinity of TH for cofactor occurs within 20 min of a single injection of a neuroloptic such as haloperidol or pimozide and lasts for at least 2 h. The action of haloperidol and other neuroleptics on CP–TH activity is abolished by either mechanical (lesion) or functional (γ-butyrolactone) interruption of nigro-neostriatal DA projections[7-9]. It is likely, therefore, that the neuronal "feedback loop" hypothesis, proposed originally by Carlsson and Lindqvist[3] to explain the regulation of nigro-neostriatial DA biosynthesis, is also relevant to the acute activation of CP–TH after systemic administration of neuroleptics. According to this theory, a neural pathway from neostriatum to substantia nigra (SN) is responsible for mediating an increase in presynaptic DA function in an effort to compensate for reduced neostriatal postsynaptic DA receptor activity.

Histochemical studies have demonstrated the presence of a GABA-containing pathway originating in the caudate nucleus with terminal projections in close proximity to soma and dendrites of nigral DA neurones[10-14]. Electro-physiological evidence indicates that the GABAergic neurones in the striatonigral pathway exert a picrotoxin-sensitive, inhibitory influence on nigral cells[15]. It is possible that a reduced release of GABA from these descending inhibitory neurones may occur as a result of neuroleptic blockade of CP–DA receptors. Activation of CP–TH could then occur as a consequence of the decreased GABAergic inhibitory control of nigral DA neurones.

The experimental verification of this hypothesis was made

possible by the use of muscimol. This isoxazol derivative, with a restricted conformation similar to that of GABA, displays bicuculline-sensitive and strychnine-insensitive inhibitory activity both *in vivo* and *in vitro*[16,17], but unlike GABA is not rapidly metabolised and thus exhibits long lasting GABA-mimetic effects *in vivo*[18]. Moreover, on the basis of experiments with a DA-sensitive adenylate cyclase, we have established that this compound is devoid of any DA-agonist or antagonist potency in homogenates of neostriatum or substantia nigra.

Muscimol was injected directly into the SN by way of chronically implanted cannulae (for detail see Figs 1 and 2). Ten minutes later haloperidol ($1.0 \, mg \, kg^{-1}$) was administered intraperitoneally and the animals were killed 30 min later for TH assay. In this experiment, TH was assayed using a concentration of cofactor ($DMPH_4$) below that required to saturate the enzyme fully.

Figure 1 shows that the haloperidol-induced increase of CP–TH activity was blocked in the caudate ipsilateral to the intranigral application of muscimol. A 50% reduction of the action of haloperidol was obtained by injection of 25 ng of muscimol into the SN; 50 ng of muscimol reliably and completely prevented the action of haloperidol on CP–TH. Muscimol (50–400 ng) alone did not alter normal CP–TH activity when measured at 20 and 40 min after intranigral injection.

The differences in the values for TH activity presented in Fig. 1 were found to reflect differences in the apparent K_m values of TH for the pteridine cofactor. Double-reciprocal plots of TH activity obtained in the presence of different cofactor concentrations and K_m values for normal, haloperidol-"activated", and muscimol-"protected" CP–TH
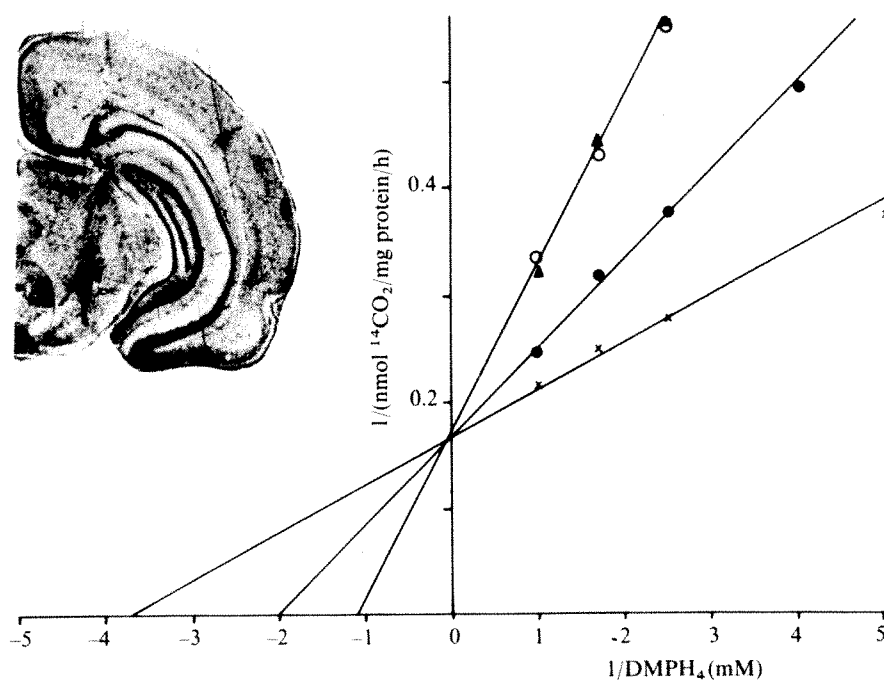
Fig. 2 Double-reciprocal plots of the initial velocity of CP–TH against various concentrations of $DMPH_4$ in the presence of 0.1 mM tyrosine. Experimental design was as described in Fig. 1. Muscimol (○, 50 ng; ●, 25 ng) or saline (×) was injected into the substantia nigra (SN). Ten minutes after intranigral injection, haloperidol (1 mg kg^{-1}) was administered intraperitoneally. The animals were killed 40 min after the intranigral injection. Control rats (▲) received saline intranigrally and vehicle intra-peritoneally. Insert: photomicrograph of coronal section of rat brain (left hemisphere) showing position of (22 g) guide-cannula tip (blackened area above the dorsal border of the SN). Injection cannula (27 g) extended 0.8 mm below guide cannula into the pars compacta region of the SN.

are shown in Fig 2. It is clear that intranigral muscimol treatment (50 ng or more) prevented the marked decrease (two- to three-fold) in apparent K_m caused by haloperidol treatment. Similar results were obtained with pimozide (1.0 mg kg^{-1}) treatment.

To determine whether the effect of intranigral muscimol was specific for the DA neurones projecting to the neostriatum, we assayed the TH activity of the nucleus accumbens of eight rats which received 25 ng of intranigral muscimol and 1.0 mg kg^{-1} haloperidol intraperitoneally. In these animals, the percentage differences between the two neostriata averaged 55% (range 40–70%), with the muscimol-treated side exhibiting lower TH activity. In contrast, the maximum bilateral difference obtained for the nucleus accumbens was less than 10%, both sides showing the usual TH activation after haloperidol. This attests to the anatomical specificity of the GABA influence on nigral DA neurones, and moreover demonstrates the degree of functional localisation for the material injected.

Data obtained from a series of rats which received intranigral bicuculline suggest that the effects we observed with muscimol are attributable to a specific GABA-mimetic action. Bicuculline methiodide (2.0 nmol per 0.4 μl over 2 min) was injected into the SN 20 min after 0.4 nmol of muscimol (50 ng). Haloperidol (1.0 mg kg^{-1} intraperitoneally) was injected 10 min before bicuculline. Control animals received muscimol and haloperidol, but with saline or strychnine (4.0 nmol per 0.4 μl over 2 min) replaced bicuculline. In all control animals, the apparent K_m of CP–TH for cofactor was 0.8 mM on the muscimol-treated side and 0.3 mM on the contralateral side. In contrast, CP–TH of all bicuculline-treated animals showed no significant bilateral differences: in every case the TH of both neostriata exhibited K_m values (0.25–0.33 mM) characteristic of haloperidol activation. Bicuculline alone had no effect on TH activity, but reversed the muscimol blockade of haloperidol-induced TH activation, and caused the dose–response relationship for muscimol to shift to the right approximately twofold. Since bicuculline (a putative GABA receptor antagonist[19]), but not strychnine (a glycine receptor antagonist[19]), reversed the action of muscimol and reinstated the effect of haloperidol on CP–TH activity, it is likely that GABA receptors mediate the action of muscimol in the SN.

Our data demonstrate that it is possible to interfere with haloperidol-induced activation of CP–TH by pharmacologically maintaining the functional effect of GABA on nigral neurones. It is unlikely that the muscimol–haloperidol interaction occurred directly in the nigra, since in other experiments we found that application of haloperidol (10^{-6}–10^{-9} M) directly into the nigra did not alter CP–TH activity. Thus the release of DA by nigral dendrites[20,21] and a postulated self-inhibitory action of DA in SN[1,22] may not be relevant to the mechanism of neuroleptic activation of CP–TH. Instead, it seems that haloperidol acts to change activity in nigral neurones indirectly, as a result of its action in the neostriatum. We suggest therefore that application of muscimol directly into the nigra, by mimicking GABA, acted to preclude the influence of a striatonigral feedback pathway. This implies that a decrease in GABA release in nigra is required for the neuroleptic-induced activation of CP–TH, a contention supported by the ability of bicuculline to reinstate TH activation after muscimol blockade.

KAREN N. GALE
A. GUIDOTTI

Laboratory of Preclinical Pharmacology,
National Institute of Mental Health,
Saint Elizabeths Hospital,
Washington, DC 20032

1 Bunney, B. S., Walters, J. K., Roth, R. H., and Aghajanian, G. K., *J. Pharmac. exp. Ther.*, 185, 560–571 (1973).
2 Carlsson, A., and Lindqvist, M., *Acta pharmac. Tox.*, 20, 140–144 (1963).
3 Neff, N., and Costa, E., *Excerpta Medica Int. Congr. Ser.*, 122, 28–34 (1966).
4 Zivkovic, B., Guidotti, A., and Costa, E., *Molec. Pharmac.*, 10, 727–735 (1974).
5 Zivkovic, B., Guidotti, A., and Costa, E., *J. Pharmac. exp.Ther.*, 194, 37–46 (1975).
6 Lovenberg, W., Bruckwick, E. A., and Hanbauer, I., *Proc. natn. Acad. Sci. U.S.A.*, 72, 2955–2958 (1975).
7 Zivkovic, B., Guidotti, A., and Costa, E., *Brain Res.*, 92, 516–521 (1975).
8 Zivkovic, B., Guidotti, A., and Costa, E., *Naunyn-Schmiedeberg's Arch. Pharmac.* 291, 193–200 (1975).
9 Zivkovic, B., Guidotti, A., and Costa, E., *J. cyclic Nucleotide Res.*, 2, 1–10 (1976).
10 McGeer, P. L., *et al.*, *Adv. behav. Biol.*, 10, 27–48 (1974).
11 Kim, J. S., Bak, I. J., Hassler, R., and Okada, Y., *Expl Brain Res.*, 14, 95–104 (1971).
12 Fonnum, I., Grafova, I., Rinvik, E., and Storm-Mathiessen, J., and Walberg, F. *Brain Res.*, 71, 77–92 (1974).
13 Okada, Y., in *GABA in Nervous System Function* (edit. by Roberts, E., Chase, T. N., and Tower, D. B.), 235–243 (Raven, New York, 1976).
14 Bak, I. J., Chio, W. B., Hassler, R., Usunoff, K. G., and Wagner, A., in *Adv. Neurol.*, 9 (edit. by Calne, D. B., Chase, T. N., and Barbeau, A.), 25–41 (Raven, New York, 1975).
15 Precht, W., and Yoshida, M., *Brain Res.*, 32, 229–233 (1971).
16 Krogsgaard-Larsen, P., Johnston, G. A. R., Curtis, D. A., Game, C. J. A., and McCulloch, R. M., *J. Neurochem.*, 25, 803–809 (1975).
17 Naik, S. R., Guidotti, A., and Costa, E., *Neuropharmacology* (in the press).
18 Waser, P. G., in *Ethnopharmacology Search For Psycho-active Drugs* (edit. by Efron, D. H., Holmstedt, B., Kline, N. S.), 419–439 (US Public Health Service Publication No. 1645, Washington, DC, 1900).
19 Curtis, D. R., and Johnston, G. A. R., *Ergebn Physiol.*, 69, 97–188 (1974).
20 Korf, J., Zieleman, M., and Westerink, B. H. C., *Nature*, 260, 257–258 (1976).
21 Geffen, L. B., Jessell, T. M., Cuello, A. C., and Iversen, L. L., *Nature*, 260, 258–260 (1976).
22 Groves, P. M., Wilson, C. J., Young, S. J., and Rebec, G. V., *Science*, 190, 522–529 (1975).

Temperature-dependent inhibition of evoked acetylcholine release in tick paralysis

IN AUSTRALIA, the tick *Ixodes holocyclus* can produce a potentially fatal, flaccid paralysis in various hosts, including man, apparently due to the secretion of a neurotoxin[1,2]. In other parts of the world other species of ticks can produce a similar syndrome[3], but it is not yet clear whether the pathogenesis of the paralysis is the same in all cases. We have found that the paralysis produced by *I. holocyclus* is due to a temperature-dependent inhibition of evoked release of acetylcholine at the neuromuscular junction.

Application of about 10 nymphal *I. holocyclus* produced paralysis in mice after 3.5–4.5 d, the exact period required depending on the rate of feeding of the ticks. Experiments were carried out on phrenic nerve–hemidiaphragm preparations removed from the most severely affected of these animals. These were mounted in a 2-ml bath and perfused continuously, at a rate of 1 ml per min, with modified Krebs' solution (120 mM NaCl, 3.5 mM KCl, 2.5 mM CaCl$_2$, 1.0 mM MgCl$_2$, 11 mM glucose, 25mM NaHCO$_3$) bubbled with 5% CO$_2$ in O$_2$. Temperature was monitored by means of a thermistor probe placed close to the recording site. The nerve was stimulated using a suction electrode and conventional intracellular techniques were used to record from the endplate region of muscle fibres. Microelectrodes were filled with 3 M KCl. Compound action potentials were recorded from the phrenic nerve using a second suction electrode. Direct stimulation of the muscle fibres was accomplished by silver wire electrodes placed on either side of the diaphragm.

At first the ability of muscles from paralysed animals to contract in response to nerve stimulation was assessed visually. At room temperature (approximately 23 °C) the muscle always contracted strongly. As the temperature was raised the response declined progressively, and at 37 °C no contraction was apparent. On subsequent reduction of the temperature the muscle response returned. The muscle contracted strongly in response to direct stimulation within this range of temperature. Similar results were obtained from all twenty preparations examined.

The effects of *d*-tubocurarine [*d*-TC] and increased magnesium chloride concentration (with compensation by reducing sodium chloride concentration) on paralysed preparations were also examined. At a concentration of 2.1 μM in the case of *d*-TC or 5 mM in the case of

magnesium chloride, the contractile response of the muscle ceased at a lower temperature than in normal solutions. At temperatures above 30 °C, no twitch occurred. In experiments on preparations from normal mice, the same concentrations of these drugs were insufficient to prevent muscle contraction.
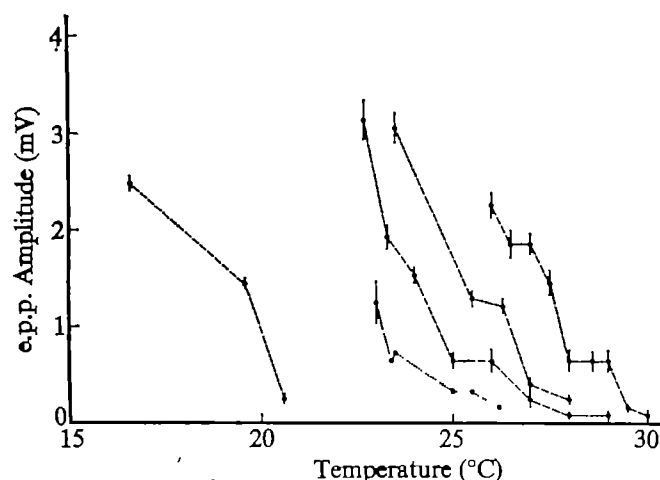
Fig. 1 Graph showing effect of temperature on e.p.p amplitude at five endplates in three different affected animals. Dashed lines indicate endplates examined in 5 mM Mg[2+]; dotted line, an endplate examined in 2.1 µM d-TC. Each point represents the mean ($\pm$ s.e.) of approximately twenty responses.

Endplate potentials (e.p.p.s) were examined in these modified solutions. In the temperature range 15–30 °C, muscle fibres could always be found in which nerve stimulation produced e.p.p.s up to 10 mV in amplitude, but none were found above 30 °C. The mean amplitudes of e.p.p.s in muscle fibres from paralysed mice were markedly temperature dependent. As the temperature was increased, the mean amplitude fell and the number of stimuli failing to produce an e.p.p. increased. In each fibre the rate of decline of e.p.p. mean amplitude appeared to be approximately the same, but the temperature range over which the decline occurred differed. This is demonstrated in Fig. 1. The effect could be reversed by lowering the temperature. The amplitude of miniature endplate potentials (m.e.p.p.s) was not altered over this temperature range (Fig. 2).

In preparations from normal mice, the effect of temperature on transmitter release was examined in 8.4 µM d-TC or 12 mM Mg. There was no decline in e.p.p. amplitude, comparable with that seen in tick paralysed preparations, over the temperature range 25–37 °C (Fig. 3).

There was no decline in the amplitude of the compound action potentials recorded from phrenic nerves removed from paralysed preparations within this range of temperature. There was, as expected[4], a small increase in amplitude as temperature was increased.

It is unlikely that the paralysis is due to a selective failure of conduction in the nerve terminals. Krnjevic and Miledi[5] have shown that this type of failure can occur, but that it results in a sudden failure of e.p.p.s and is more likely to occur at low temperatures.

Another possibility is that the toxin produces, a temperature-dependent, graded reduction of action potential amplitude in the nerve terminals. This possibility has not been entirely eliminated. It seems unlikely, however, that the amplitude of the nerve terminal action potential could be reduced sufficiently to abolish transmitter release without a detectable effect on that of the nerve trunk.
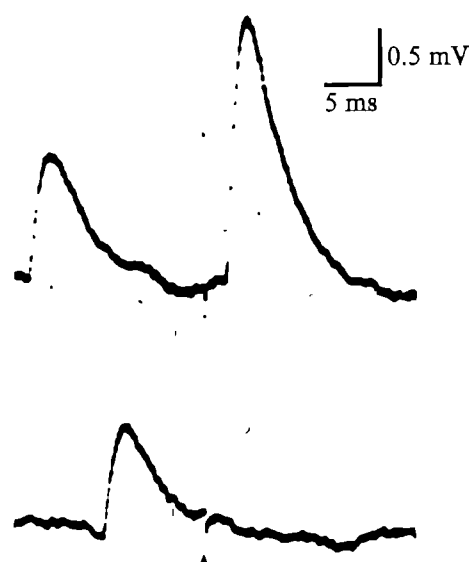
Fig. 2 Endplate potentials and m.e.p.p.s from the same affected endplate at 26.5 °C (upper trace) and 28.5 °C (lower trace). Arrow indicates stimulus artefact.

Other workers[6,7] have suggested that the paralysis produced by the tick *Dermacentor andersoni* is associated with a reduction in amplitude of the compound action potential recorded from peripheral nerves *in vivo*. We found no evidence of this in our experiments and it may be that the modes of action of the toxins from the two species of tick differ. A central action for the toxin of *D. andersoni* has also been postulated[8,9]. Our observations, however, suggest that most aspects of the motor paralysis produced by *I. holocyclus* can be explained on the basis of action at the neuromuscular junction.
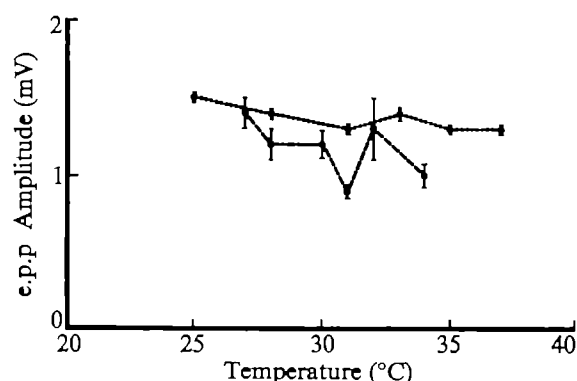
Fig. 3 Graph showing effect of temperature on e.p.p. amplitude in preparations from normal mice. Dashed line indicates an endplate examined in 12 mM Mg[2+]; dotted line an endplate examined in 8.4 µM d-TC. Each point represents the mean ($\pm$ s.e.) of approximately twenty responses.

The data presented here suggest that the toxin secreted by *I. holocyclus* has a direct, temperature-sensitive action on the excitation–secretion coupling mechanism, which results in an inhibition of transmitter release at the neuromuscular junction. This inhibition seems to be due to a reduction in quantal content, rather than quantal size, as the latter would be expected to result in m.e.p.p.s of reduced amplitude. Because spontaneous release continued when evoked release was blocked, it seems likely that the toxin does not interfere with the actual release mechanism, but rather with some intermediate step between depolarisation of the terminal membrane and release. An attractive

hypothesis is that the toxin blocks the influx of calcium ions which seems to be essential before evoked release can occur[10,11]. This remains to be investigated.

B. J. COOPER

Department of Veterinary Pathology,
University of Sydney,
New South Wales 2006

I. SPENCE*

Department of Physiology and Pharmacology,
University of New South Wales,
Kensington, New South Wales 2003

1 Clunies Ross, I., *J. Council sci. ind. Res.*, **8**, 8–13 (1935).
2 Kaire, G. H., *Toxicon*, **4**, 91–97 (1966).
3 Gregson, J. D., *Can. Dept Agric. Mono.*, No 9, 10–12 (1973).
4 Gasser, H. S., *Am. J. Physiol.*, **97**, 254–270 (1931).
5 Krnjevic, K., and Miledi, R., *J. Physiol., Lond.*, **149**, 1–22 (1959).
6 Emmons, P., and McLennan, H., *J. exp. Biol.*, **37**, 355–362 (1960).
7 Murnaghan, M. F., and MacConaill, M., *Irish J. med. Sci.*, sixth series, **502**, 473–477 (1967).
8 Esplin, P. W., Philip, C. B., and Hughes, L. E., *Science*, **132**, 958–959 (1960).
9 Emmons, P., and McLennan, H., *Nature*, **183**, 474–475 (1959).
10 Katz, B., and Miledi, R., *Proc. R. Soc.*, **B167**, 23–38 (1967).
11 Katz, B., and Miledi, R., *J. Physiol., Lond.*, **207**, 789–801 (1970).

Mesolimbic dopaminergic neurones in the rotational model of nigrostriatal function

AT least some of the symptoms of Parkinson's disease are believed to result from the degeneration of nigrostriatal dopaminergic neurones which normally innervate the striatum. Parkinsonian patients have an abnormally low concentration of dopamine in this region of the basal ganglia[1,2], and parkinsonian symptoms can be alleviated by the dopamine precursor L-dopa[3]. A quantitative animal model which has proved useful in assessing the potential therapeutic value of drugs in the treatment of parkinsonism was described some years ago by Ungerstedt and Arbuthnott[4]. Rats, in which the nigrostriatal pathway has been destroyed unilaterally by 6-hydroxydopamine (6-OHDA), were shown to rotate towards the lesioned side when injected with drugs, such as amphetamine, which release dopamine from neurones in the brain, and towards the unlesioned side when injected with L-dopa and dopamine agonists[5,6]. This contralateral rotation has been attributed to supersensitivity of the denervated striatal dopamine receptors[5,6].

We have investigated now whether another system of dopaminergic neurones, the mesolimbic system is involved in the drug-induced rotational behaviour. This system, comprised of dopaminergic neurones which originate in the A10 group of cell bodies and which innervate the nucleus accumbens and olfactory tubercle[7], has recently been shown to be important in the locomotor behaviour produced by various drugs[8-15]. In a recent study of the possible role of mesolimbic dopamine neurones in rotational behaviour it was shown that amphetamine and the dopamine agonist apomorphine caused no rotation in rats with unilateral 6-OHDA-induced destruction of mesolimbic dopamine neurones[16]. However, it remained possible that bilateral destruction of the mesolimbic dopamine system could affect the drug-induced rotation of rats with unilateral lesions of the nigrostriatal pathway. We have now found this to be the case.

Adult male Sprague–Dawley albino rats received an intrastriatal injection of 6-OHDA into the right caudate nucleus and during the same operation bilateral injections of 6-OHDA or vehicle into the nucleus accumbens. 6-OHDA injections were made stereotaxically while the rats were anaesthetised with Equithesin (3 ml kg⁻¹). Stereotaxic coordinates for the nucleus accumbens were $+3.4$, 1.7, 7.2 according to the atlas of Pellegrino and Cushman[17], and for the caudate nucleus were $+2.0$, 3.0, 5.5. Injections were made through a 30-gauge stainless steel cannula at the rate of 1 μl min⁻¹. Eight micrograms (calculated as base) of 6-hydroxydopamine hydrobromide (Sigma), dissolved in 0.9% saline containing ascorbic acid (1 mg ml⁻¹) was injected in a volume of 2 μl. Vehicle injections were identical except for the omission of 6-OHDA from the solution. At various times following these lesions the rotation induced by intraperitoneal injections of (+)-amphetamine sulphate or the dopamine agonist apomorphine (hydrochloride) was measured. The rats were placed in hemispherical white plastic bowls and the number of turns per min recorded every 10 min by direct observation. Fifty-five to sixty days after the operation the rats were killed by decapitation and their brains removed for the measurement of regional catecholamine content. The brains were chilled on an ice-cold glass plate and the olfactory tubercles, nuclei accumbens, right and left striata, and neocortex were dissected as described previously[18,19]. Tissues were homogenised in 0.1 N perchloric acid containing 0.1 mg ml⁻¹ EDTA and assayed for noradrenaline and dopamine by a radioenzymatic method[20] modified from that of Cuello *et al.*[21] The significance of statistical differences was calculated using Student's *t* test.

Fig. 1 shows the dopamine concentration in the nucleus accumbens and striata of the two groups of animals. In both groups the dopamine content of the right striatum was reduced to 35% of that of the left ($P<0.01$). In addition, the 6-OHDA lesion of the nucleus accumbens reduced the dopamine content of the nucleus accumbens to 25% of control ($P<0.01$) and that of the olfactory tubercle to 41% (6.7 ± 0.6 against 2.7 ± 0.4 μg g⁻¹, $P<0.01$). Neocortical noradrenaline content also was reduced to 49% (0.35 ± 0.01
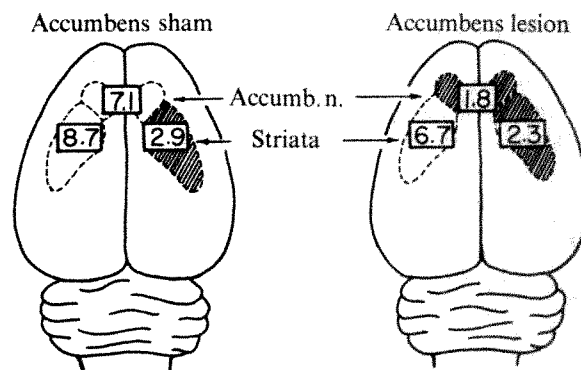
Fig. 1 Diagrammatic representation of the dopamine concentration (μg g⁻¹) in the nucleus accumbens and left and right striata of the two groups of rats (7 rats per group) used in the circling experiments. Shaded areas indicate 6-OHDA (8 μg in 2 μl) was injected into the region.

against 0.17 ± 0.01 μg g⁻¹, $P<0.01$) of control by the nucleus accumbens lesion.

Figure 2 shows the responses of the two groups of animals to intraperitoneal injections of amphetamine (5 mg kg⁻¹) and apomorphine (0.5 mg kg⁻¹). The effect of the 6-OHDA lesion of the nucleus accumbens was to block the ipsilateral rotation produced by amphetamine, and to enhance the contralateral rotation produced by apomorphine. Although the 6-OHDA lesion of the nucleus accumbens destroyed both mesolimbic dopamine neurons and noradrenergic neurones innervating the neocortex we believe the behavioural effects

of the lesion result from the destruction of mesolimbic dopamine neurones as we have recently obtained similar results when rats were treated with desipramine before the 6-OHDA injection. This pretreatment protects noradrenergic neurones from destruction by 6-OHDA[13,22]. It therefore seems that activity in the mesolimbic dopaminergic system is as important as an imbalance of striatal dopaminergic activity for the expression of drug-induced rotational behaviour. When amphetamine-induced release of dopamine from mesolimbic neurones is prevented by their destruction amphetamine-induced rotation is blocked. The rotation produced by the directly acting dopamine agonist apomorphine, however, is enhanced, presumably because the denervated mesolimbic dopamine receptors become supersensitive.

These findings have implications for the use of rotational models in assessing the effects of drugs as agonists or antagonists at dopamine receptors. First it is clear that the drug-induced rotation of rats with unilateral 6-OHDA-induced destruction of nigrostriatal neurones is not a pure model of striatal dopaminergic mechanisms. Rather it seems that rotational behaviour depends on an imbalance of striatal dopaminergic activity plus activity in the mesolimbic dopamine system. The predominant effect of activity in the mesolimbic system seems to be on the rate of circling rather than its direction. In contrast to the nigrostriatal system there appears to be no directional influence of mesolimbic dopaminergic activity. Rats with unilateral 6-OHDA-induced destruction of mesolimbic dopamine neurones do not rotate in response to amphetamine or apomorphine[16]. Neither do unilateral injections of dopamine agonists into the nucleus accumbens cause any rotation[15].

As a tool for screening potential dopamine agonists and antagonists the rotation model remains useful. It would, however, be wrong to assume that the effects of dopamine agonists and antagonists in this preparation were due solely to their effects on striatal dopaminergic mechanisms. For example, drug-induced rotation would be blocked by drugs which blocked only mesolimbic or only striatal dopamine receptors. This point may be of more than theoretical interest as there is some evidence for differences between mesolimbic and striatal dopamine receptors[24]. The rotation produced by dopamine agonists can clearly be affected by the actions of the drug on mesolimbic dopamine receptors,

but our results do not predict whether a drug must act as an agonist at both mesolimbic and nigrostriatal dopamine receptors to produce rotation. It is conceivable, for example, that activation of nigrostriatal dopamine receptors by the drug and release of endogenous dopamine in the mesolimbic system would be sufficient to cause rotation. However, after injections of dopamine or dopaminergic agonists into the striatum the major effect seems to be postural asymmetry rather than a quantitative rotational response[23]. Activation of mesolimbic dopamine receptors by dopamine agonists may therefore be a necessary condition for these drugs to produce rotation.

Our results show that a drug-induced behavioural expression of nigrostriatal output, rotation, is markedly modified by the concomitant activity at mesolimbic dopamine receptors. Probably, therefore, activity at mesolimbic dopamine receptors can modify the effects of nigrostriatal activity in the undrugged animal, and the mesolimbic and nigrostriatal systems may interact physiologically in the control of motor behaviour. The connections which have been described between the nucleus accumbens and the caudate nucleus[25] and substantia nigra[25-27] might provide an anatomical basis for mesolimbic-nigrostriatal interactions such as those which we have described here. Finally, it is worthy of comment that this and previous work[8-15] emphasises a role of the mesolimbic dopamine system in motor function. Its pathology in motor dysfunction has very recently been reported (O. Hornykiewicz, unpublished).

P. H. KELLY
K. E. MOORE

Department of Pharmacology,
Michigan State University,
East Lansing, Michigan 48824

1 Ehringer, H., and Hornykiewicz, O., *Klin. Wschr.*, 38, 1236–1239 (1960).
2 Bernheimer, H., Birkmayer, W., and Hornykiewicz, O., *Klin. Wschr.*, 41, 465–469 (1963).
3 Cotzias, G. C., van Woert, M. H., and Schiffer, L. M., *New Engl. J. Med.*, 276, 374–379 (1967).
4 Ungerstedt, U., and Arbuthnott, G. W., *Brain Res.*, 24, 485–493 (1970).
5 Ungerstedt, U., *Acta physiol. scand.*, Suppl., 367, 69–93 (1971).
6 Von Voigtlander, P. F., and Moore, K. E., *Neuropharmacology*, 12, 451–462 (1973).
7 Ungerstedt, U., *Acta physiol. scand.*, Suppl. 367, 1–48 (1971).
8 Pijnenburg, A. J. J., Woodruff, G. N., and Van Rossum, J. M., *Brain Res.*, 59, 289–302 (1973).
9 Pijnenburg, A. J. J., Honig, W. M. M., and Van Rossum, J. M., *Psychopharmacologia*, 41, 87–95 (1975).
10 Kelly, P. H., Seviour, P. W., and Iversen, S. D., *Brain Res.*, 94, 507–522 (1975).
11 Kelly, P. H., Miller, R. J., and Neumeyer, J. L., *Eur. J. Pharmac.*, 35, 85–92 (1976).
12 Kelly, P. H., and Iversen, L. L., *Psychopharmacologia*, 45, 221–224 (1975).
13 Iversen, S. D., and Kelly, P. H., in *Chemical Tools in Catecholamine Research*, 1 (edit. by Jonsson, G., Malmfors, T., and Sachs, C.), 327–333 (North-Holland–American Elsevier, New York, 1975).
14 Woodruff, G. N., Kelly, P. H., and Elkhawad, A. O., *Psychopharmacology*, 47, 195–198 (1976).
15 Elkhawad, A. O., and Woodruff, G. N., *Br. J. Pharmac.*, 54, 107–114 (1975).
16 Kelly, P. H., *Brain Res.*, 100, 163–169 (1975).
17 Pellegrino, L. J., and Cushman, A. J., *A Stereotaxic Atlas of the Rat Brain* (Appleton-Century-Crofts, New York, 1968).
18 Glowinski, J., and Iversen, L. L., *J. Neurochem.*, 13, 655–669 (1966).
19 Horn, A. S., Cuello, A. C., and Miller, R. J., *J. Neurochem.*, 22, 265–270 (1974).
20 Moore, K. E., and Phillipson, O. T., *J. Neurochem.*, 25, 289–294 (1975).
21 Cuello, A. C., Hiley, R., and Iversen, L. L., *J. Neurochem.*, 21, 1337–1340 (1973).
22 Breese, G. R., and Traylor, T. D., *Br. J. Pharmac.*, 42, 88–99 (1971).
23 Ungerstedt, U., Butcher, L. L., Butcher, S. G., Andén, N-E., and Fuxe, K., *Brain Res.*, 14, 461–471 (1969).
24 Fuxe, K., et al., *J. Pharmac.*, (*Paris*), 6, 117–129 (1975).
25 Powell, E. W., and Leman, R. B., *Brain Res.*, 105, 389–403 (1976).
26 Rioch, D. M., *J. comp. Neurol.*, 53, 319–388 (1931).
27 Smith, O. C., *J. comp. Neurol.*, 51, 65–127 (1930).

Fig. 2 Turning produced by (+)-amphetamine (*a*) (5 mg kg^{-1}, i.p.) or apomorphine (*b*) (0.5 mg kg^{-1} intraperitoneally) in rats with unilateral 6-hydroxydopamine (6-OHDA) lesions of the caudate nucleus with (●) or without (○) additional bilateral 6-OHDA lesions of the nucleus accumbens. Positive values represent turning towards the lesioned caudate while negative values denote the opposite direction. Mean ±s.e.m. for groups of 7 rats.

Fusion of human erythrocyte ghosts promoted by the combined action of calcium and phosphate ions

CALCIUM ions are important in almost all membrane fusion systems[1,2]. Murayama and Okada[3] showed that Sendai virus-induced fusion of Ehrlich ascites cells required Ca^{2+}. When

Ehrlich ascites cells were treated with Sendai virus in the presence of EDTA, cells agglutinated and lysed but did not fuse[3,4]. Similarly, Ca[2+] was required for the fusion of erythrocytes stimulated by chemicals such as lysolecithin[5] or glyceryl monooleate[6]. Ca[2+] plus the ionophore A23187 (ref. 7) or Ca[2+] at high pH (ref. 8) promoted the fusion of chicken erythrocytes. Fusion of phosphatidylserine-rich liposomes was absolutely dependent on the presence of Ca[2+9]. We describe here the fusion of human erythrocyte ghosts promoted by the combined action of Ca[2+] and phosphate buffer. Human erythrocyte. ghosts, because they lack cytoplasm and are easily filled with small or large molecules[10,11] are useful for the investigation of membrane fusion.

calcium phosphate formed in the experimental conditions. In Tricine or Tris buffer plus Ca[2+] neither agglutination nor fusion was observed. To obtain fusion, it was essential to incubate the ghosts in phosphate buffer and then to add the Ca[2+]. If the Ca[2+] and the phosphate buffer were preincubated for 20 min at 37 °C before addition of erythrocyte ghosts, calcium phosphate precipitated. When ghosts were added to such a precipitate, there was some agglutination followed by lysis but no fusion. Calcium phosphate has been implicated in the facilitation of extracellular DNA .uptake in eukaryotic cells[13].

Table 2 shows the effects of changing the pH of the phosphate buffer on the agglutination and fusion of erythrocyte ghosts in

Table 1 Agglutination and fusion of human erythrocyte ghosts in presence of bivalent cations and various buffers

Experiment	Buffer	Salt	Concentration (mM)	Agglutination	Fusion
(1)	Phosphate	None			
	Phosphate	CaCl₂	2	+++	+++
	Phosphate	BaCl₂	2	+	+
	Phosphate	BaCl₂	4	++	++
	Phosphate	MnCl₂	2	++++	—
	Phosphate	ZnCl₂	2	++++	—
	Phosphate	MgCl₂	2	—	—
(2)	Phosphate	None		—	—
	Phosphate	CaCl₂	2	+++	+++
	Arsenate	None		—	—
	Arsenate	CaCl₂	2	—	Trace
	Tricine–NaOH	None		—	—
	Tricine–NaOH	CaCl₂	2	—	—
	Tris–HCl	None		—	—
	Tris–HCl	CaCl₂	2	—	—

Out-dated human blood type O was obtained from the blood bank of Hadassah Hospital, Jerusalem. Blood aged from 4–8 weeks was used (stored blood) In experiment (1) stored human erythrocytes were washed three times with a solution containing 150 mM KCl and 10 mM sodium phosphate buffer, pH 7.4 (isotonic solution) as described before[11]. To prepare ghosts, 2.5 ml of 25% (v/v) erythrocytes in the isotonic medium were dialysed for 2 h against 2 l of hypotonic medium containing 40 mM KCl and 10 mM sodium phosphate buffer, pH 7.4. Ghosts were resealed by adding 0.125 ml of a solution containing 2.2 M KCl and 0.02 M MgSO₄ and incubating them at 37 °C for 15 min. The ghosts were then washed three times in the isotonic solution. Erythrocyte ghosts were fused as follows. Reaction mixtures contained a final concentration of ghosts equivalent to 1.25% (v/v) of the original erythrocytes in a solution containing 120 mM KCl, 30 mM NaCl and 10 mM sodium phosphate buffer, pH 7.4. The experiment was initiated by adding bivalent metal ions as shown in the table and incubating the suspensions at 37 °C for 20 min. In experiment (2) erythrocyte ghosts were prepared as described for experiment (1) except that the phosphate buffer in the isotonic and hypotonic solutions was replaced where indicated by 10 mM Tricine–NaOH buffer, 10 mM Tris–HCl buffer or 10 mM sodium arsenate buffer, pH 7.4. Ghosts were fused as described for experiment (1). Agglutination and fusion were estimated by phase contrast microscopy. Agglutination: +, clump of 5–10 cells; ++, clumps of 10–20 cells; +++, clumps of 50–100 cells; ++++, big clumps of 100 cells or more. Fusion: +, fusion of two to four cells, and 5–10% of the ghosts fused; ++, polyghosts derived from 5–10 cells, and 10–20% of ghosts fused; +++, large polyghosts derived from 10–20 cells and 60–80% of ghosts fused

Table 1 shows the effects of various bivalent cations and different buffers on the agglutination and fusion of erythrocyte ghosts. Ca[2+] plus phosphate buffer was the most effective combination for promoting fusion. Ba[2+] (Table 1) and Sr[2+] (not shown) with phosphate buffer caused a moderate amount of fusion, and Ca[2+] and arsenate buffer caused a very small degree of fusion. Fusion was maximal with Ca[2+] and phosphate buffer after 15–20 min of incubation at 37 °C but the 'poly-ghosts' formed were unstable and disintegrated after a further 10–15 min. Ca[2+], Mn[2+], Zn[2+] and Ba[2+] agglutinated ghosts in the presence of phosphate buffer (Table 1). Zn[2+] and Mn[2+] in phosphate buffer did not promote fusion although they were active agglutinating agents. There seems to be a close correlation between the formation of a precipitate with phosphate ions and the agglutinating effect of bivalent cations. Mg[2+], the only cation tested which does not form a precipitate with phosphate, did not agglutinate ghosts. All bivalent metal ions which formed precipitates with phosphate caused agglutination. Fusion thus seems to be distinct from agglutination, and more specific in its bivalent metal ion requirement. Furthermore, in other experiments not shown here, ghosts agglutinated with Ca[2+] plus phosphate buffer disaggregated when titrated with acetic acid. The amount of acetic acid required for disaggregation was identical to that required to dissolve the precipitate of

the presence of Ca[2+]. Erythrocyte ghosts were agglutinated from pH 7.0 upwards; and fusion was maximal at pH 7.5. There was a direct correlation between the amount of calcium phosphate precipitated, which increased with pH, and the extent of agglutination.

Table 2 Effect of pH on agglutination and fusion of human erythrocyte ghosts by Ca[2+] plus phosphate

pH	Agglutination	Fusion	Remarks
5	—	—	
5.5	—	—	
6	—	—	
6.5	+	—	
7	++	+	Fusion after 60 min
7.5	+++	+++	Fusion after 20 min
8	++++	++	Fused cells disinte-
8.5	++++	+	grated shortly after
9	++++	+	fusion

Erythrocyte ghosts were prepared as described in Table 1 for experiment (1). Reaction mixtures contained a final concentration of ghosts equivalent to 1.5% (v/v) erythrocytes in a solution containing 120 mM KCl, 30 mM NaCl, and 10 mM sodium phosphate buffer, at the pH indicated. The experiment was initiated by adding CaCl₂ at a final concentration of 2 mM and the suspensions were incubated immediately at 37 °C for 20 min. Agglutination and fusion were estimated as described in Table 1.

Table 3 Effect of internal ATP on fusion of human erythrocytes or erythrocyte ghosts

Experiment	Cells	Treatment	Addition during fusion	Agglutination	Fusion
(1)	Fresh human erythrocytes	Stored at 4 °C	None	—	—
			2 mM CaCl₂	+ + + +	—
		Incubated at 37 °C with glucose, adenine, inosine and phosphate	None	—	—
			2 mM CaCl₂	+ + + +	—
		Incubated at 37 °C with NaF (ATP-depleted)	None	—	—
			2 mM CaCl₂	+ + + +	+ +
(2)	Human erythrocyte ghosts		None	—	—
			2 mM CaCl₂	+ + +	+ + +
	ATP-containing human erythrocyte ghosts		None	—	—
			2 mM CaCl₂	+ + +	—

In experiment (1), fresh human blood type O was used. ATP-depleted blood was prepared as described before[26]. The erythrocytes were washed in a solution containing 135 mM KCl, 5.4 mM NaCl, 0.8 mM MgSO₄ and 20 mM Tricine–NaOH, pH 7.4. Blood, 2% (v/v) was incubated overnight at 37 °C in the presence of 10 mM NaF for depletion of ATP or in the presence of 5 mM glucose, 10 mM adenine, 2 mM inosine and 5 mM sodium phosphate as a control. The samples were washed three times in a solution containing 120 mM KCl, 30 mM NaCl and 20 mM Tricine–NaOH, pH 7.4. Blood not incubated overnight was washed in the same solution. The three kinds of erythrocytes were suspended in the above solution at a concentration of 1.25% (v/v) in the presence of 2 mM CaCl₂ and 5 mM glucose (glucose was omitted in the depleted blood) for 30 min at 37 °C. The samples were then washed three times in a solution containing 120 mM KCl, 30 mM NaCl and 10 mM sodium phosphate buffer, pH 7.4, and suspended at a concentration of 2.5% (v/v). For fusion, CaCl₂ at a final concentration of 2 mM was added to each of the erythrocyte suspensions and they were transferred immediately to 37 °C for 20 min with slow shaking. Agglutination: + + + +, big clumps of more than 100 cells. Fusion: + +, polyerythrocytes made of 10–30 cells and 20–30% of the cells were fused. In experiment (2) stored human erythrocytes type O + were washed three times in a solution containing 160 mM KCl and 20 mM Tricine–NaOH, pH 7.4. Ghosts were prepared as follows: ATP was trapped in human erythrocyte ghosts by dialysis of 1.25 ml of 25% (v/v) of erythrocytes in the presence or absence of 5 mM ATP for 2 h, against 1 l of a hypotonic medium containing 40 mM KCl and 10 mM Tricine–NaOH, pH 7.4 at 0 °C. To 1 ml of each system of ghosts 0.050 ml of a solution containing 2.2 M KCl and 0.02 M MgSO₄ was added. The ghosts were resealed for 5 min at 37 °C then washed in a solution containing 150 mM KCl and 10 mM sodium phosphate buffer, pH 7.4 and suspended at a concentration equivalent to 25% (v/v) of the original erythrocytes. Fusion was achieved as described in Table 1, experiment (1). The cells were incubated at 37 °C for 30 min. Agglutination and fusion were estimated as described in Table 1, experiment (1).

Fresh human erythrocytes incubated with Ca^{2+} and phosphate buffer at pH 7.4 agglutinated but did not fuse (Table 3). It seemed possible that the failure of intact erythrocytes to fuse was attributable to the low concentration of internal Ca^{2+} maintained by their Ca^{2+} pump. Since depletion of ATP inhibits the Ca^{2+} pump which expels Ca^{2+} from the erythrocyte cytoplasm[13], the effect of ATP depletion on fusion was tested. Table 3 shows that ATP-depleted erythrocytes fused readily in phosphate buffer plus Ca^{2+}. The polyerythrocytes formed were unstable and disintegrated about 10 min after fusion. The inclusion of ATP in erythrocyte ghosts causes internal Ca^{2+} to be pumped out of them[14], so it was of interest to test the fusion of ghosts containing trapped ATP. Erythrocyte ghosts containing ATP agglutinated, but did not fuse, in conditions where ghosts without ATP fused freely (Table 3). These observations could be explained by assuming that fusion of erythrocytes or their ghosts requires intracellular Ca^{2+}. In earlier experiments, Ahkong et al.[7] showed that chicken erythrocytes, pretreated with neuraminidase, fused when Ca^{2+} was introduced to the cytoplasmic side of the membrane with cation ionophore A23187.

We have achieved fusion of human erythrocyte ghosts without the aid of virus, using Ca^{2+} and phosphate ions at a physiological pH and temperature. Previously, virus-induced fusion of human erythrocyte ghosts was observed only after gradual haemolysis in the presence of bovine serum albumin[15] or after drastic haemolysis followed by treatment with sulphydryl group-blocking agents[16].

Addition of Ca^{2+} to the medium causes the aggregation of intramembrane particles of human erythrocyte ghosts[17]. Presumably the Ca^{2+} penetrates to the cytoplasmic side of the membrane where it causes polymerisation of spectrin and consequent intramembrane particle aggregation[17]. It has been suggested that fusion of cells involves the rearrangement of intramembrane particles[18]. Because of the apparent interrelationships of Ca^{2+}, intramembrane particle rearrangement and

fusion, we used freeze-etching to investigate the distribution of intramembrane particles in ghosts fused with Ca^{2+} and phosphate ions. Figure 1a shows a section of erythrocyte ghosts fused with Ca^{2+} and phosphate ions. A polyghost agglutinated with unfused ghosts can be seen. Figure 1b shows a freeze-fracture picture of two cells whose membranes are in close juxtaposition. Most of the area of the inner phase of the membrane in both cells contains crowded intramembrane particles. The area of close contact of the two cells (arrow) is interesting in that it is almost devoid of intramembrane particles, and, there seems to be an interdigitation of membrane projections. It seems likely that this is an area of contact just before fusion.

Figure 1c is a freeze-fracture picture from a ghost preparation after fusion. The fracture face of the inner phase of the membrane consists of two areas studded with tightly packed intramembrane particles connected by a folded smooth area almost devoid of particles. It is tempting to suggest that the smooth area is a region of fusion joining two cells. No such smooth areas were detected in unfused controls. Since intramembrane particles are probably membrane proteins or glycoproteins, and fusion of cells has been suggested to occur between areas of exposed phospholipids[19], intramembrane particles might have been removed from regions in which fusion occurs. Previous work in our laboratory showed that on clustering of intramembrane particles, phospholipids are exposed and become susceptible to attack by chemical agents and phospholipases[20,21].

There are strong indications that Ca^{2+} is involved in certain biological processes involving membrane fusion such as secretion of macromolecules[22], endocytosis[23], fusion of myoblasts[24], and fusion of Golgi vesicles[25]. It is possible that fusion in these cases occurs by a similar mechanism to that suggested here.
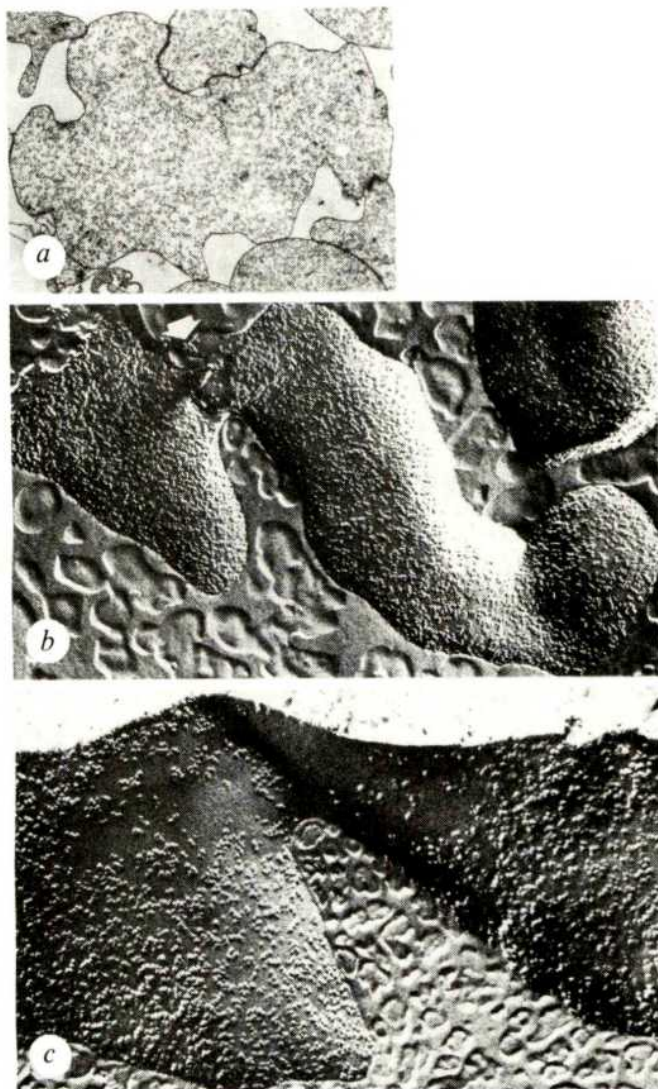
Fig. 1 Electron micrographs of human erythrocyte ghosts in the process of fusing. Human erythrocyte ghosts were prepared and fused as described in the legend to Table 1, experiment (1). The bivalent metal used was CaCl₂ (2 mM). Samples were incubated for 25 min at 37 °C for fusion. Preparation of sections for electron microscopy (*a*) was as described before[27]. Freeze-fractured cells were prepared essentially as described before[20] with the following modifications. Erythrocyte ghosts were fixed immediately after fusion by addition of glutaraldehyde (Ladd Research Industries) to a final concentration of 1 % (v/v) and centrifuged at 12,000*g* for 5 min. *a*, Electron micrograph of a polyghost tightly agglutinated with unfused ghosts (× 4,800). *b*, Freeze-fracture picture of two cells tightly agglutinated before fusion. Interdigitated region of contact (arrow) is poor in intramembrane particles (× 38,280). *c*, Freeze-fracture picture of fused ghosts. Note folded smooth area separating surfaces with tightly packed intramembrane particles. The smooth area is presumed to be a region of fusion between two ghosts (× 61,200).

NEHAMA ZAKAI
RICHARD G. KULKA
ABRAHAM LOYTER

*Department of Biological Chemistry,
Hebrew University of Jerusalem,
Jerusalem, Israel*

1 Okada, Y., in *Current Topics in Microbiology and Immunology*, 48 (edit. by Arber, W., and Braun, W.), 102–128 (Springer, Berlin, 1969).
2 Poste, G., and Allison, A. C., *Biochim. biophys. Acta*, 300, 421–465 (1973).
3 Murayama, F., and Okada, Y., *Biken's J.*, 8, 103–105 (1965).
4 Okada, Y., and Murayama, F., *Expl Cell Res.*, 44, 527–551 (1966).
5 Poole, A. R., Howell, J. I., and Lucy, J. A., *Nature*, 227, 810–813 (1970).
6 Ahkong, Q. F., Fisher, D., Tampion, W., and Lucy, J. A., *Biochem. J.*, 136, 147–155 (1973).
7 Ahkong, Q. F., Tampion, W., and Lucy, J. A., *Nature*, 256, 208–209 (1975).
8 Toister, Z., and Loyter, A., *Biochim. biophys. Acta*, 241, 719–724 (1971).
9 Papahadjopoulos, D., Poste, G., Schaeffer, B. E., and Vail, W. J., *Biochim. biophys. Acta*, 352, 10–28 (1974).
10 Seeman, P., *J Cell Biol.*, 32, 55–70 (1967).
11 Loyter, A., Zakai, N., and Kulka, R. G., *J. Cell Biol.*, 66, 292–304 (1975).
12 Graham, F. L., and Van der Eb, A. J., *Virology*, 52, 456–467 (1973).
13 Romero, P. J., and Whittam, R., *J. Physiol., Lond.*, 214, 481–507 (1971).
14 Olson, E. J., and Cazort, R. J., *J. gen. Physiol.*, 53, 311–322 (1969).
15 Peretz, H., Toister, Z., Laster, Y., and Loyter, A., *J. Cell Biol.*, 63, 1–11 (1974).
16 Michaeli, D., Lalazar, A., and Loyter, A., *Israel J. med. Sci.*, 11, 1185 (1975).
17 Elgsaeter, A., Shotton, D. M., and Branton, D., *Biochim. biophys. Acta*, 426, 101–122 (1976).
18 Satir, R., in *Symp. Soc. exp. Biol.*, 28 (edit. by Sleigh, M. A., and Jennings, D. H.), 399–418 (Cambridge University Press, London, 1974).
19 Lucy, J. A., *Nature*, 227, 814–817 (1970).
20 Gazitt, Y., Loyter, A., Reichler, Y., and Ohad, I., *Biochim. biophys. Acta*, 419, 479–492 (1976).
21 Gazitt, Y., Ohad, I., and Loyter, A., *Biochim. biophys. Acta* (in the press).
22 Rubin, R. P., *Pharmac. Rev.*, 22, 389–428 (1970).
23 Allison, A. C., and Davies, P., in *Sym. Soc. exp. Biol.*, 28 (edit. by Sleigh, M. A., and Jennings, D. H.), 419–446 (Cambridge University Press, London, 1974).
24 Shainberg, A., Yagil, G., and Yaffe, D., *Expl Cell Res.*, 58, 163–167 (1969).
25 Gratzl, M., and Dahl, G., *FEBS Lett.*, 62, 142–145 (1976).
26 Gazitt, Y., Ohad, I., and Loyter, A., *Biochim. biophys. Acta*, 382, 65–72 (1975).
27 Toister, Z., and Loyter, A., *J. biol. Chem.*, 248, 422–432 (1973).

T lymphocytes with promiscuous cytotoxicity

CYTOTOXIC T lymphocytes generated during a unidirectional mixed lymphocyte culture (MLC) lyse target cells which have the antigenic phenotype of the allogeneic stimulating cells. The cytotoxic effect is restricted to cells that bear the same major histocompatibility antigens (H–2 antigens in mice) as the stimulating cells[1]. H–2 restriction of cytotoxicity is also seen when immune T lymphocytes react *in vitro* to cells bearing viral[2], chemical[3], or minor histocompatibility antigens[4] or following autosensitisation against unmodified fibroblasts[5]. Thus, H–2 restriction of cytotoxic lymphocytes has been observed in many circumstances. We now report that cultures of normal spleen cells generate T lymphocytes that damage target cells regardless of their H–2 phenotype.

To investigate the influence of alloantigens on the generation of cytotoxic T lymphocytes, we cultured normal mouse spleen in the presence or absence of allogeneic stimulator spleen cells or fibroblasts. After 5 d culture, the responder lymphocytes were tested against various fibroblast monolayer targets using a modification of a standard microcytotoxicity assay. In this system the number of adherent fibroblasts remaining after 48 h of incubation with the cultured spleen cells was measured by uptake of radioactive ⁵¹Cr. As seen in Table 1a, C57L (H–2ᵇ) spleen cells, after incubation with irradiated CBA stimulator cells (H–2ᵏ), caused fourfold greater cytotoxicity of B10.Br (H–2ᵏ) than C57BL/6 (H–2ᵇ) fibroblasts. This demonstrated the expected relative H–2 restriction of T-cell cytotoxicity which results from an allogeneic MLC. The presence of allogeneic fibroblasts also induced H–2-restricted cytotoxicity (Table 1b).

In contrast, spleen cells cultured alone, without allogeneic stimulator cells, showed a similar high degree of cytotoxicity against target fibroblasts without H–2 restriction (Table 1c). In each of 20 experiments, normal spleen cells cultured for 5 d developed promiscuous cytotoxicity that did not discriminate among the H–2 phenotypes of target fibroblasts.

To identify the effector cell in the population of cultured spleen cells, the influence of anti-θ treatment[6] was tested with the results shown in Table 2. The marked decrease of cytotoxic activity after anti-θ + C′ treatment suggests that a T lymphocyte was necessary for the expression of promiscuous cytotoxicity. This conclusion was supported by the finding that the cytotoxic effect was not decreased by filtering the effector cells through nylon wool. Thus, promiscuous cytotoxicity, like H–2 restricted cytotoxicity[7], seems to be a function of T lymphocytes. Cell-free medium

in which effector lymphocytes had been cultured for 48 h showed no cytotoxicity against target fibroblasts when compared with fresh medium. Therefore there was no evidence for a nonspecific cytotoxic factor produced by the lymphocytes.

Experiments were carried out to identify the requirements for induction of promiscuous cytotoxicity. Fresh spleen cells or spleen cells cultured for up to 48 h did not produce this effect. Cells cultured for 3 d had 50% of maximal cytotoxic activity. Peak activity was usually

Table 1 Induction of H-2-restricted or promiscuous cytotoxicity by culture of normal spleen cells

Spleen cells Strain	H–2	Allogeneic stimulator cells Strain	H–2	Target fibroblasts Strain	H–2	% Cytotoxicity†
a C57L	b	CBA spleen cells	k	C57L	b	27
				B10.BR	k	98
b B10	b	B10A fibro- blasts	a	B10	b	3
				B10.A	a	83
B10.A	a	B10 fibro- blasts	b	B10.A	a	8
				B10	b	84
				B10.D2	d	7
				B10.BR	k	5
c C57L	b	None		C57L	b	75
				B10.BR	k	67
B10	b	None		B10	b	92
				B10.D2	d	91
				B10.BR	k	89
				B10.A	a	96
B10.A	a	None		B10.A	a	93
				B10	b	94
				B10.D2	d	94
				B10.BR	k	94
NZB	d	None		NZB	d	78
				B10.A	a	73
C57BL/6	b	None		C57BL/6	b	76
				NZB	d	81
BALB/c	d	None		DBA/2	d	90
				C57BL/6	b	92

*Suspension of spleen cells of 6–8-week-old male mice were prepared by forcing fragments of spleen through a tantalum gauze mesh. The spleen cells were suspended in medium consisting of RPMI 1640 (Grand Island, Gibco) plus 10% foetal calf serum (FCS) (heat inactivated at 56% for 30 min; Gibco), 5×10^{-5} M 2-mercaptoethanol, sodium pyruvate 1 mM (Microbiological Associates), non-essential amino acids 0.1 mM (Microbiological Associates) and penicillin (5,000 units ml^{-1}) and streptomycin (5,000 μg ml^{-1}) (RPMI + 10% FCS). The spleen cells were washed, centrifuged and resuspended at a concentration of 7×10^6 viable nucleated cells in 2 ml of RPMI + 10% FCS. The spleen cells were incubated for 5 d in 16-mm tissue culture wells (Costar) in a volume of 2 ml per well at 37 °C in moist air plus 10% CO_2. After induction the spleen cells were collected from the culture wells by repeated pipetting of the medium, washed by centrifugation and tested for cytotoxicity. Some induction cultures contained allogeneic stimulator cells which were either 3.5×10^6 irradiated (2,000 r.) spleen cells (*a*), or monolayers of 10^5 fibroblasts (*b*). The fibroblasts were prepared from 13–17-d-old mouse embryos as described[15].

†Cytotoxicity was assayed in quadruplicate by incubating test spleen cells with target fibroblasts at effector–target ratios of either 50:1 or 100:1. Control cultures contained fresh uncultured spleen cells of the same genotypes as the test spleen cells. One million spleen cells in 2 ml RPMI + 10% FCS were added to either 10^4 or 2×10^4 target fibroblasts which had been cultured for one day in 16-mm culture wells. After 44 h, the spleen cells and detached fibroblasts were aspirated and the adherent fibroblasts washed four times with RPMI × 10% FCS. To measure the relative numbers of adherent fibroblasts, we incubated each culture well for 40 min with 3 μCi ^{51}Cr in 0.2 ml of 0.3 M sucrose. The wells were then washed twice with RPMI + 10% FCS and aspirated dry. Uptake of ^{51}Cr was determined by digesting the fibroblasts with 1 ml 0.1 N NaOH. After 30 min the contents of each well were transferred to plastic tubes and radioactivity was measured in gamma counter. Percentage cytotoxicity was computed as (1−mean c.p.m. (target with test spleen cells/target with control fresh spleen cells) × 100). The standard deviations were always less than 10% of the mean c.p.m.

attained after 4–5 d of culture and persisted undiminished for up to 9 d.

Adherent cells such as macrophages are involved in the induction of cytotoxic T lymphocytes by allogeneic cells[8]. The role of an adherent population in the induction and effector phases of promiscuous cytotoxicity was investigated with the results shown in Table 3. Spleen cells were plated on Petri dishes for 1 h at 37 °C. The non-adherent cells were removed, cultured alone for 5 d and then assayed for cytotoxicity. The non-adherent cells had a markedly decreased capacity to generate cytotoxic cells when compared with an unseparated population. Conversely, absorption of spleen cells on Petri dishes after 5 d of culture did not reduce their capacity to mediate cytotoxicity. Fractionation of spleen cells on nylon wool columns before or after *in vitro* culture confirmed the requirement for an adherent cell in the induction of cytotoxicity, but not in the effector phase.

The inducer of promiscuous cytotoxicity is unknown. Foetal calf serum in the culture medium might have contained mitogen-like substances capable of inducing polyclonal differentiation of effector lymphocytes[9]. It is also

Table 2 Effect of anti-θ treatment on promiscuous cytotoxicity

Treatment of C57BL/6 cytotoxic spleen cells	% Cytotoxicity of C57BL/6 fibroblast targets
None	94
C′	91
Anti-θ	93
Anti-θ + C′	33

Promiscuous cytotoxicity was induced and tested as in Table 1. After induction, the spleen cells were untreated, or treated with complement (C) anti-θ globulin or anti-θ globulin + C′ and tested for cytotoxicity against syngeneic fibroblasts (see Table 1). Anti-θ globulin (Cohn Fr. II of AKR anti C3Hθ serum) was added to spleen cells (30×10^6–50×10^6) at a dilution of 1:10 in 0.3 ml and incubated at room temperature for 45 min. C′ (fresh guinea pig serum absorbed with C57BL/6 spleen cells for one hour at 0 °C) was used at a final dilution of 1:9 in 1 ml and incubated at 37 °C for 1 h. Some spleen cells were treated sequentially with anti-θ globulin followed by C′.

possible that the foetal calf serum was specifically immunogenic during induction. Its association with the fibroblasts in the target cultures could therefore have triggered cytotoxicity to the fibroblasts[10]. This is unlikely, however, since we found that adding either allogeneic or syngeneic fibroblasts to the inducing cultures led to H-2 restriction of the resulting cytotoxicity (our unpublished results).

Regardless of the underlying mechanism, it is clear that *in vitro* culture of spleen cells can lead to the generation of T lymphocytes with relatively unrestricted cytotoxicity. How then does addition of allogeneic cells to the culture abort promiscuous cytotoxicity? It is possible that the relevant alloantigens may favour the selective proliferation of clones. Alternatively, the induction of specific cytotoxicity by allogeneic cells may be accompanied by the active suppression of unselected clones.

The emergence of nonspecific suppressor activity following *in vitro* culture of normal spleen cells has been observed by Hodes and Hathcock[11] and by Burns *et al.*[12]. "Precultured" spleen cells could not be induced into cytotoxic effectors in an MLC. Such cells also suppressed the responses of fresh spleen cells to allogeneic or modified syngeneic targets[11]. The induction of antibody production *in vitro* was also inhibited by spleen cells cultured in a similar manner[12]. This suppression was mediated by a T lymphocyte and had similar kinetics and cellular requirements for induction as the promiscuous cytotoxicity reported here.

What is the relationship between promiscuous cytotoxicity and the reports of suppressor activity that developed spontaneously in cultures of mouse spleen cells? These

Table 3 Effect of removal of adherent cells on promiscuous cytotoxicity

Absorption of BALB/c spleen cells on Petri dishes	% Cytotoxicity of NZB fibroblast targets
None	76
Before induction	18
After induction	70

Promiscuous cytotoxicity was induced and tested as in Table 1. Some of the spleen cells were absorbed before induction by placing 20×10^6 cells in 3 ml RPMI + 10% FCS in 60-mm tissue culture dishes (Falcon) and incubating at 37 °C for 1 h. Non-adherent cells were removed by repeated pipetting, cultured for 5 d and tested for cytotoxicity. Some spleen cells were cultured for 5 d and then absorbed as above. The non-adherent cells were then tested for cytotoxicity.

phenomena conceivably result from activities of distinct subclasses of T lymphocytes: cytotoxic and suppressor cells. Alternatively, these phenomena may reflect the dual function of a single subclass of lymphocytes. In other words, these different *in vitro* measurements may be manifestations of the same biological function. Suppression of an MLC or antibody production[11,12] could result from indiscriminate damage to the reacting lymphocytes. And the effect we observed may be yet another manifestation of suppressor cell function. Identity of suppressor and cytotoxic T lymphocytes is also supported by the finding that the cells mediating these activities share the same Ly-2,3 T-cell marker[13,14].

C.S. is supported by a fellowship from the MRC of Canada. This work was supported by grants from the USPHS.

CHAIM SHUSTIK
IRUN R. COHEN*
ROBERT S. SCHWARTZ
ELIZABETH LATHAM-GRIFFIN

Hematology Service,
New England Medical Center Hospital,
Boston, Massachusetts 02111

SAMUEL D. WAKSAL

Immunology Branch,
National Cancer Institute,
National Institutes of Health,
Bethesda, Maryland 20014

1 Hayry, P., and Defendi, V., *Science*, 168, 133–135 (1970).
2 Zinkernagel, R. M., and Doherty, P. C., *J. exp. Med.*, 141, 1427–1436 (1975).
3 Shearer, G. M., Rehn, T. G., and Garvarino, C. A., *J. exp. Med.*, 141, 1348–1364 (1975).
4 Gordon, R. D., Simpson, E., and Samelson, L. E., *J. exp. Med.*, 142, 1108–1120 (1975.)
5 Ilfeld, D., Carnaud, C., and Klein, I., *Immunogenetics*, 2, 231–240 (1975).
6 Raff, M. C., *Immunology*, 19, 637–650 (1970).
7 Wagner, H., Harris, A. W., and Feldmann, M., *Cell. Immun.*, 4, 39–50 (1972).
8 Wagner, H., Feldmann, M., Boyle, W., and Schrader, J. W., *J. exp. Med.*, 136, 331–343 (1972).
9 Anderson, J., Sjoberg, O., and Moller, G., *Transplant. Rev.*, 11, 131–177 (1972).
10 Forni, G., and Green, T., *J. Immun.*, 166, 1561–1565 (1976).
11 Hodes, R. J., and Hathcock, R. S., *J. Immun.*, 116, 167–177 (1976).
12 Burns, F. D., Marrack, P. C., Kappler, J. W., and Janeway, C. A., *J. Immun.*, 114, 1345–1347 (1975).
13 Jandinski, J., Cantor, H., Tadakuma, T., Peavy, D. L., and Pierce, C. W., *J. exp. Med.*, 143, 1382–1390 (1976).
14 Cantor, H., and Boyse, E. A., *J. exp. Med.*, 141, 1376–1389 (1975).
15 Wekerle, H., Lonai, P., and Feldmann, M., *Proc. natn. Acad. Sci. U.S.A.*, 69 1620–1624 (1972).

New principle for the analysis of chemical carcinogenesis

THE development of cancer following exposure to chemical carcinogens or to various forms of irradiation is almost invariably slow and prolonged. Although the process can be initiated by a brief exposure to a carcinogenic stimulus, there is no evidence that target cells so altered are cancer cells. Rather, there is abundant indirect evidence from many systems that what is induced is an altered cell or cell population from which malignant neoplasia can gradually develop or evolve[1,2]. Neoplastic development therefore resembles a chain reaction, triggered by exposure to a carcinogen, in which the links are new populations with altered organisational, structural and biochemical properties. These slowly proliferative new lesions are characteristically focal in distribution, implying that only a small proportion of the original target cell population in any organ or tissue participates. It is not known what the critical property (or properties) is that makes initiated cells so important in carcinogens and the failure to understand and manipulate this early step has been a major impediment to its analysis.

We have developed a new approach to the study of chemical carcinogenesis which seems to offer the first quantitative assay for populations of initiated cells, as well
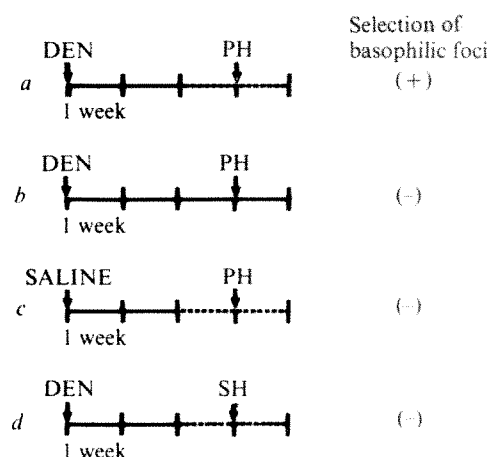
Fig. 1 Schematic representation of the assay procedure (*a*) for initiated cells in liver induced by a carcinogen. *b*, *c*, *d*, Essential controls. DEN, 200 mg per kg body weight intraperitoneally; —, 1 week of basal diet; – – –, 1 week basal diet plus 0.02% AAF; PH, partial hepatectomy; SH, sham hepatectomy.

as the possibility of a sequential analysis of some of the early putative preneoplastic and premalignant hepatocytes for liver cancer. The principle is functional rather than morphological or biochemical and depends on the increasing evidence that the selective growth of focal precancerous liver cells seems to result from their relative resistance to the cytotoxic action of hepatocarcinogens, coupled with the creation of a local environment for their proliferation.

All known hepatocarcinogens inhibit regeneration of liver cells following partial hepatectomy and at higher doses may lead to liver-cell death. The focal, slowly proliferative cell populations called hyperplastic nodules, that occur in the liver before the appearance of cancer, all acquire new properties relating to the metabolism and toxic effects of hepatocarcinogens and other hepatotoxins that require metabolic activation. These include resistance to hepatotoxin-induced cell death[3], decreased uptake and/or activation of some carcinogens[3], a uniform decrease in several components of microsomal mixed function monooxygenase system including cytochrome $P450$, aryl hydrocarbon hydroxylase and aminopyrine demethylase[4,5] and also the capacity for proliferation following partial hepatectomy during carcinogen administration[6,7].

In the light of these considerations, we wondered

whether an early, or even first, change induced by an hepatocarcinogen might be the induction of an altered metabolic pattern in some liver cells (similar to that seen in hyperplastic nodules), such that these cells would have a potential growth advantage over the majority of original hepatocytes in a cytotoxic environment. If this were the case, it should be possible to detect the altered cells soon after initiation by creating an intense stimulus for proliferation in the presence of a toxic growth inhibitor requiring metabolic activation. Initiated cells with a decreased capacity for activation might proliferate rapidly, whereas cells unaltered in this respect would activate the carcinogen and thereby generate a derivative which would inhibit their proliferation. On the basis of this theory, we devised an assay system for such resistant cells (Fig. 1a).

The assay system consists of three components: an initiator, for example, diethylnitrosamine (DEN); a selective growth inhibitor, 2-acetylaminofluorene (2-AAF), and a generalised potent growth stimulus, in this case, partial hepatectomy (PH). The major carcinogen so far studied has been DEN, since it can induce liver cancer in the rat slowly with a single oral, intravenous, or intraperitoneal administration[8,9]. DEN was given intraperitoneally to Fischer-344 rats at a dose of 200 mg kg^{-1}. Following a 2-week period of recovery from the initial cell damage, the animals were fed a standard basal diet (24% protein) containing 0.02% 2-AAF for 1 week and were then subjected to 67% partial hepatectomy. With only 2-AAF and partial hepatectomy, there was no hepatocyte proliferation, and no labelling of hepatocyte nuclei with ^{3}H-thymidine as observed autoradiographically. In animals pretreated with DEN, however, there were multiple focal islands of proliferating intensely basophilic hepatocytes (Fig. 2) distributed randomly throughout the liver which started to appear about 30 h after operation. They grew rapidly, so that within 7–10 d they are visible grossly as small, greyish-white nodules about 1 mm in diameter. No such foci were seen in animals treated with saline in place of DEN, given 2-AAF without PH, subjected to PH without concomitant exposure to 2-AAF or subjected to sham operation. Although subsequent studies have shown that feeding the 2-AAF diet for about 3 weeks before PH is sufficient to induce a few basophilic foci, they were rarely seen in the present experiments without pretreatment with DEN.
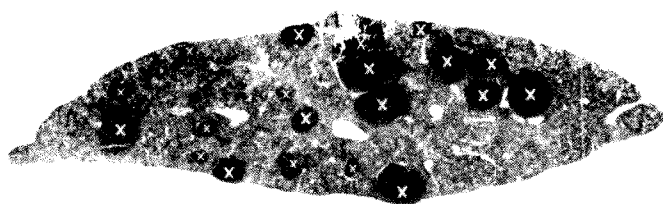
The number of foci induced by DEN varies with the dose. Figure 3 shows the relative 'focus-forming ability' over a dose range of 20 to 200 mg kg^{-1}. The foci could be counted either by low power examination of histological sections (Fig. 2) or at 7–10 d by counting the foci in a gross liver section, the sections of liver having been taken from a standard site. Foci similar to those seen with DEN have been induced by every hepatocarcinogen studied to date, including dimethylnitrosamine, aflatoxin B1, and N-hydroxy-2-AAF.

The initiated foci can be labelled with ^{3}H-thymidine without labelling the surrounding original hepatocytes and

in this way, the fate of the hepatocytes in the foci can be followed. If the dietary 2-AAF is discontinued following partial hepatectomy, selective proliferation of the foci stops as the majority of original hepatocytes recover the ability to proliferate. By manipulating the conditions, it thus becomes possible to control the number of cell proliferations and to study their role in the development of cancer.

Following termination of the specialised regimen at 7 d after operation, the foci continue to grow, so that by 8 weeks the liver is composed of many hyperplastic nodules indistinguishable from those described previously[10,11]. If the animal is not subjected to the 'selection pressure' used (2-AAF plus PH), the latent or potential foci persist for at least 10 weeks without any significant change in size or number. Thus, the initiated hepatocytes induced by DEN have not yet acquired the property of autonomous growth but can be induced to grow in an appropriate environment. Also, we have obtained no evidence for their recognition by the host or for their removal or repair. The later development of some autonomous growth of such cells represents another step in the sequence before the ultimate development of malignant neoplastic cells.
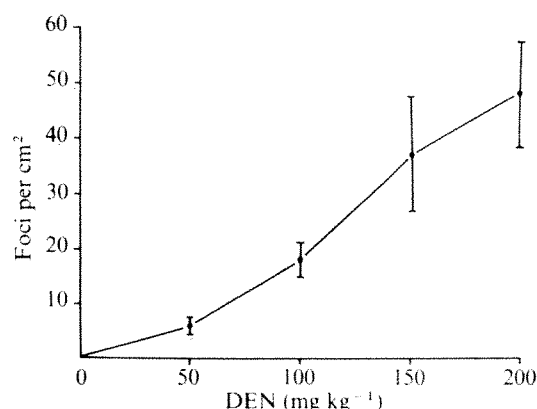
Fig. 3 Induction of foci of initiated liver cells as a function of the dose of diethylnitrosamine (DEN). The number of foci was counted in sections from each lobe by low power microscopy and is expressed per cm^2 of cross-sectional area (mean ± s.e.m.).

Scherer and Emmelot detected histochemically and counted microscopic foci of enzyme-deficient hepatocytes induced with DEN[12]. These foci may prove to 'be the initiated populations selected in the present system.

The interval between the administration of carcinogen and the assay is being shortened so that the kinetics of induction of the initiated foci can be established clearly with different carcinogens. Another desirable modification would be the use of a non-carcinogen as the selective growth inhibitor. Even in its present form, however, the assay is useful for the study of several aspects of chemical carcinogenesis, including the possible role of DNA synthesis and of repair of specific DNA damage in initiation, the persistence of initiated cells, the sequential analysis of the properties of hepatocytes acquired after initiation, the possible functional theme underlying the development to liver cancer and the clarification of some confusing aspects in the use of liver hepatoma formation in mice as an assay for carcinogens.

The hypothesis of differential cytotoxicity as an important theme for chemical carcinogenesis is not new[4,10,13–17], but it has never been subjected to critical tests. On the basis of our results, it is tempting to consider that the acquisition of resistance to some of the toxic effects of an added carcinogen or, in the case of a single exposure to a carcinogen, to toxic xenobiotic compounds in the natural

diet or absorbed bacterial products in the intestinal tract, may be the major motive force during liver carcinogenesis until the precancerous cells can grow autonomously. Consistent with this view of liver carcinogenesis are the findings of Peraino and colleagues on the acceleration of liver carcinogenesis by phenobarbital or DDT in animals exposed for only 18 d to 2-AAF (refs 18, 19). An underlying hypothesis for liver carcinogenesis is now developing, which is amenable to experimental tests. Whether the principle of selective cytotoxicity can be extended to the development of cancer in other organs or tissues remains to be explored.

DENNIS SOLT
EMMANUEL FARBER

Department of Pathology,
University of Toronto,
Toronto, Ontario, M5G 1L5, Canada

1 Foulds, L., *Neoplastic Development*, 1, ch. 3 (Academic, London, 1969).
2 Farber, E., *Cancer Res.*, 33, 2537–2550 (1973).
3 Farber, E., Parker, S., and Gruenstein, M., *Cancer Res.* (in the press).
4 Gravela, E., Feo, F., Canuto, R. A., Garcea, R., and Gabriel, L., *Cancer Res.*, 35, 3041–3047 (1975).
5 Cameron, R., Sweeney, G. D., Jones, K., Lee, G., and Farber, E., *Cancer Res.* (in the press).
6 Becker, F. F., and Klein, K. M., *Cancer Res.*, 31, 169–173 (1971).
7 Kitagawa, T., *Gann*, 62, 217–224 (1971).
8 Druckrey, H., Stemhoff, D., Preussman, R., and Ivankovic, S., *Naturwissenschaften*, 50, 735 (1963).
9 Craddock, V. M., *Chem.-Biol. Interact.*, 10, 313–321 (1975).
10 Farber, E., *Methods in Cancer Res.*, 7, 345–375 (1973).
11 Sasaki, T., and Yoshida, I., *Virchows Arch. path. Anat. Physiol.*, 295, 175–300 (1935).
12 Scherer, E., and Emmelot, P., *Eur. J. Cancer*, 11, 689–696 (1975).
13 Haddow, A., *Acta Un. Int. Cancr.*, 3, 342–352 (1938).
14 Laws, J. D., *Br. J. Cancer*, 13, 669–674 (1959).
15 Vasiliev, J. M., Guelstein, V. I., *J. natn. Cancer Inst.*, 31, 1123–1143 (1963).
16 Diamond, L., *Prog. exp. Tumour Res.*, 11, 364–383 (1969).
17 Kitagawa, T., Michalopoulis, G., and Pitot, H. C., *Cancer Res.*, 35, 3682–3692 (1975).
18 Peraino, C., Fry, R. J. M., Staffeldt, E., and Kisieleski, W. E., *Cancer Res.*, 33, 2701–2705 (1973).
19 Peraino, C., Fry, R. J. M., Staffeldt, E., and Christopher, J. P., *Cancer Res.*, 35, 2884–2890 (1975).

Ammonia assimilation in lupin nodules

ROBERTSON *et al.*[1,2] proposed that plant enzymes rather than bacteroid enzymes are physiologically important in the assimilation of the ammonia produced after nitrogen reduction in the bacteroids of leguminous nodules. They demonstrated that during the development of the nodule, glutamine synthetase (EC 6.3.1.2, equation (1)) and glutamate synthase (EC 2.6.1.53, equation (2)) activity increased considerably in the plant fraction but not in the bacteroid fraction of the nodule, parallel with the induction of nitrogenase and leghaemoglobin.

$$\text{glutamate} + NH_4^+ + ATP \longrightarrow \text{glutamine} + ADP + P_i \quad (1)$$

$$\text{glutamine} + \text{2-oxoglutarate} + NADH$$
$$\longrightarrow \text{2 glutamate} + NAD^+ \quad (2)$$
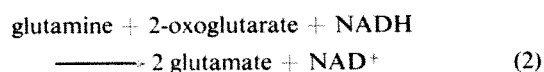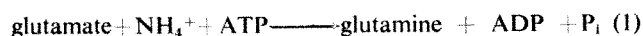
But, the predominant amino acid transported in the xylem of lupins and many other legumes is asparagine[2–4]. Robertson *et al.*[2] suggested therefore that asparagine was most probably synthesised in the plant fraction of the nodule by a glutamine-dependent asparagine synthetase[5,6] (EC 6.3.5.4, equation (3)) rather than in the bacteroids, which might utilise for asparagine biosynthesis an ammonium-dependent asparagine synthetase

(EC 6.3.1.1, equation (4)), which has been reported to occur in other prokaryotes[7].

$$\text{aspartate} + \text{glutamine} + ATP \longrightarrow \text{asparagine}$$
$$+ \text{glutamate} + AMP + PP_i \quad (3)$$

$$\text{aspartate} + NH_4^+ + ATP \longrightarrow \text{asparagine} + AMP + PP_i \quad (4)$$
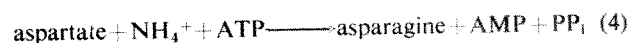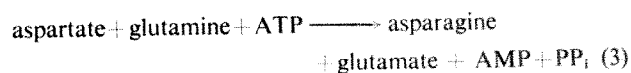
We now report the formation of asparagine by a cell-free extract of the plant fraction of lupin nodules in the presence of ATP, aspartate, glutamine and Mg^{2+} (Table 1). Apparent K_m values for aspartate and glutamine were determined to be 3.6 mM and 0.26 mM, respectively. Ammonia could replace glutamine as the source of the amide nitrogen; but a fourfold greater concentration of ammonia gave only a rate of synthesis half that observed with glutamine (Table 1). Asparagine synthetase activity could not be detected unless the nodules were crushed in the presence of a sulphydryl protecting agent such as dithiothreitol. Furthermore, the enzyme was found to be stabilised by inclusion of 5 mM glutamine or 6 mM 5-diazo-4-oxo-L-norvaline (DONV), a glutamine analogue in the buffer. These properties of the asparagine synthetase from the plant fraction of *Lupinus angustifolius* nodules are in good agreement with the properties of the enzyme recently purified from etiolated *Lupinus luteus* cotyledons[8].

Asparagine synthetase activity in the plant fraction of the developing nodules was first detected 13 d after the seedlings were inoculated with rhizobia (Fig. 1). The rapid increase in activity seen after day 13 is similar to that already reported for the plant glutamine synthetase (ref. 1 and repeated here Fig. 1), plant glutamate synthase[2], nitrogenase[1] and leghaemoglobin[1]. No asparagine synthetase activity could be detected in the plant fraction of 11 or 12-d-old nodules (Fig. 1). But these fractions contained an asparaginase activity whereas the plant fractions from 13-and 24-d-old nodules did not (Fig. 1). Attempts to detect asparagine synthetase activity in the plant fraction of 11-and 12-d-old nodules in which the asparaginase was inhibited 80–90% with DONV (a known inhibitor of bacterial asparaginases[9]) were unsuccessful (Fig. 1).

Table 1 Synthesis of asparagine by extracts of the plant fraction of lupin nodules

Incubation mixture	Asparagine synthetase activity (nmol asparagine per min per mg protein)
Complete system	2.33
−glutamine	0.30
−glutamine + 13.2 mM NH_4Cl	1.23
−Mg^{2+}	ND
−ATP	ND
Boiled enzyme	ND

Lupins (*Lupinus angustifolius* L. var. Uniwhite) were inoculated with *Rhizobium lupini* (NZP 2257) and grown in a controlled environment cabinet[2]. Nodules were collected from 25-d-old plants, crushed in 0.5 M sucrose, 50 mM K phosphate buffer, *p*H 7.8, containing 5 mM dithiothreitol and separated into a bacteroid and plant fraction as previously described[2]. Asparagine synthetase activity was measured at 30 °C by the formation of ^{14}C-asparagine in a total volume of 300 µl containing plant fraction (200 µl), 4 mM L-4-^{14}C-aspartate (0.83 µCi µmol⁻¹), 3.3 mM $MgCl_2$, 3.3 mM L-glutamine and 4 mM ATP. Samples of 25 µl were removed during 15 min of incubation and applied to a Whatman 3 MM paper. Asparagine was separated from aspartate by electrophoresis at *p*H 5.0 and the radioactivity in each amino acid was determined[1]. Identification of the product as asparagine was confirmed by electrophoresis and chromatography using two procedures: (1) electrophoresis in formic acid–water (4% v/v) adjusted to *p*H 2.2 with pyridine (50 V cm⁻¹ for 2 h) followed by descending chromatography in *n*-butanol–*n*-butyl acetate–acetic acid–water (19:1:5:25 v/v) and (2) electrophoresis in 0.04 M sodium acetate, *p*H 5.0, (30 V cm⁻¹ for 30 min) followed by descending chromatography in methanol–pyridine–water (80:4:20 v/v). Furthermore, the ^{14}C-asparagine eluted from these chromatograms was converted to ^{14}C-aspartic acid by hydrolysis in 2 M HCl at 105 °C for 16 h.
ND, not detectable, less than 0.2 nmol per min per mg of protein.

The decrease in asparaginase activity in the plant fraction at days 11 and 12 coupled with the complete absence of this enzyme on day 13 (Fig. 1) was particularly interesting. This plant asparaginase was different from the bacteroid or free-living rhizobial enzyme in terms of (1) electrophoretic mobility on 7.5% polyacrylamide gels relative to bromophenol blue (0.47 for the plant enzyme and 0.27 for the bacteroid or rhizobial enzyme) and (2) apparent K_m for asparagine (7 mM for the plant enzyme and 6 μM for the rhizobial enzyme). In view of the fact that the plants were grown in a nitrogen-free medium, it seemed likely that until the bacteroids developed nitrogen-fixing ability, the plant asparaginase had a role in supplying the developing nodule tissue and bacteroids with ammonia by hydrolysing asparagine transported from the cotyledons. Once nitrogen fixation had been established 11–12 d after inoculation the plant asparaginase would no longer be required and activity would be lost (Fig. 2). An analogous role for a plant asparaginase in supplying rapidly developing legume seed tissue with C and N has been proposed by Atkins *et al.*[10].

So far we have been unable to detect any ammonium- or glutamine-dependent asparagine synthetase activity in either the bacteroid fraction of the nodules or in cell-free extracts of the free-living rhizobia grown in a broth or minimal medium. Attempts were made to inhibit or remove an asparaginase activity from these extracts before assaying for asparagine synthetase, but in no case could any synthetase activity be detected.

Our studies reported here and elsewhere[1,2] have led us to suggest that plant enzymes rather than bacteroid enzymes are involved in the assimilation of the ammonia produced from nitrogen reduction in the bacteroid. The first step in this assimilation (Fig. 2) is proposed to be the synthesis of glutamine by the plant glutamine synthetase. It has been estimated previously[2] that the levels of this enzyme found in the plant fraction of the nodule would be able to incorporate into glutamine more than twice the amount of ammonia actually produced by the bacteroids in 18-d-old lupin nodules. On the other hand,

Fig. 2 Proposed pathway for the assimilation of the ammonia produced after nitrogen reduction in the bacteroids into asparagine by enzymes located in the plant fraction of lupin nodules.

less than 3% of this ammonia could be assimilated by the bacteroid enzyme[2], a conclusion recently confirmed by the work of Kurz *et al.*[11] and Brown and Dilworth[12]. The high level of this plant glutamine synthetase, reported here in lupin and in other legume nodules[12], coupled with the low K_m for NH_4^+ (0.02 mM)[13] would provide an efficient mechanism for rapidly removing any ammonia passing out of the bacteroid into the plant. In the presence of this enzyme it is unlikely therefore that the plant asparagine synthetase would be able to incorporate ammonia directly into asparagine since the K_m for NH_4^+ for the plant asparagine synthetase is 3–5 mM (refs 5 and 6).

After the incorporation of ammonia into glutamine in the plant fraction of the nodule tissue, the amide nitrogen of glutamine would according to our model be transferred through glutamate using glutamate synthase and a transaminase to oxaloacetate to give the α-amino of aspartate (Fig. 2). The synthesis of asparagine by the glutamine-dependent asparagine synthetase would then be accomplished using this aspartate and a second amide nitrogen from glutamine (Fig. 2). The net result of these reactions would be the synthesis of asparagine from oxaloacetate and 2 mol of ammonium requiring NADH and 3 ATP.

The role of asparagine as the major amino acid carrier of fixed nitrogen is supported by an examination of the amino acid

Fig. 1 Levels of asparagine synthetase, asparaginase and glutamine synthetase during nodule development in lupin. Nodules were collected from plants collected at intervals after inoculation and the plant enzyme fraction was prepared as in Table 1. Asparagine synthetase activity was assayed as for Table 1 (●) or after DONV treatment[2] (○). Glutamine synthetase levels (■) were determined as previously described[1]. Asparaginase activity (□) was determined radiochemically using the procedure outlined for glutaminase assays[1], except that ^{14}C-U-asparagine was used as substrate instead of glutamine.

Table 2 Amino acid analysis of xylem sap from 11- and 18-d-old lupin plants

Amino acid*	μmol amino acid per ml per sap	
	11 d	18 d
Asparagine†	0.77	2.35
Aspartate	0.32	0.54
Glutamine†	0.23	0.41
Glutamate	0.33	0.36
Threonine	0.15	0.40

Xylem sap was collected using a pressure chamber[11] and stored under liquid nitrogen. Amino acid analyses were carried out using a JEOL 6A-H analyser.

*The level of all other amino acids was less than 0.1 μmol per ml of sap.

†Determined as the increase in aspartate or glutamate on hydrolysis for 4 h at 105 °C in 1 M HCl.

levels in the xylem sap of 11- and 18-d-old plants (Table 2). Although asparagine is the predominant amino acid transported in both cases, it is the asparagine level which increases and not that of glutamine as the capability of the nodule to fix nitrogen increases. Transpiration rates determined using 18-d-old plants were such that the plants would be transporting asparagine at a rate of 2 µmol per h per plant. Furthermore, we can estimate from the activity of asparagine synthetase reported here, that the nodules could synthesise up to 0.8 µmol asparagine per h per plant, which is in reasonable agreement with the rate of transportation of this amide considering the instability of asparagine synthetase *in vitro*.

In conclusion, we have proposed a pathway by which ammonia produced by the bacteroids is assimilated by the plant fraction of lupin nodules. Presumably the bacteroid enzyme systems provide the ATP and electrons for N_2-reduction whereas the plant enzyme systems provide the additional NADH and ATP required for the incorporation of ammonia into asparagine. Further interest will undoubtedly lie in elucidating the mechanisms regulating the synthesis of nitrogenase in the bacteroids, and leghaemoglobin, glutamine synthetase, glutamate synthase and asparagine synthetase in the plant fraction of the nodule. The regulation of the plant asparaginase activity and the reason why the bacteroids overproduce ammonia for plant growth beyond the level required for bacterial growth are also intriguing questions demanding further study.

D. BARRY SCOTT
KEVIN J. F. FARNDEN

Department of Biochemistry,
University of Otago,
Dunedin, New Zealand

JOHN G. ROBERTSON

Applied Biochemistry Division, DSIR
Palmerston North, New Zealand

1 Robertson, J. G., Farnden, K. J. F., Warburton, M. P., and Banks, J. M., *Aust. J. Pl. Physiol.*, **2**, 265–272 (1975).
2 Robertson, J. G., Warburton, M. P., and Farnden, K. J. F., *FEBS Lett.*, **55**, 33–37 (1975).
3 Pate, J. S., Gunning, B. E. S., and Briarty, L. G., *Planta*, **85**, 11–34 (1969).
4 Streeter, J. G., *Agron. J.*, **64**, 311–314 (1972).
5 Rognes, S. E., *FEBS Lett.*, **10**, 62–66 (1970).
6 Streeter, J. G., *Archs Biochem. Biophys.*, **157**, 613–624 (1973).
7 Cedar, H., and Schwartz, J. H., *J. biol. Chem.*, **244**, 4112–4121 (1969).
8 Rognes, S. E., *Phytochemistry*, **14**, 1975–1982 (1975).
9 Jackson, R. C., and Handschumacher, R. E., *Biochemistry*, **9**, 3585–3590 (1970).
10 Atkins, C. A., Pate, J. S., and Sharkey, P. J., *Pl. Physiol., Lancaster*, **56**, 807–812 (1975).
11 Kurz, W. G. W., Rokosh, D. A., and LaRue, T. A., *Can. J. Microbiol.*, **21**, 1009–1012 (1975).
12 Brown, C. M., and Dilworth, M. J., *J. gen. Microbiol.*, **86**, 39–48 (1975).
13 O'Neal, D., and Joy, K. W., *Pl. Physiol., Lancaster*, **54**, 773–779 (1974).
14 Scholander, P. F., Hammel, H. T., Hemmingsen, E. A., and Bradstreet, E. D., *Proc. natn. Acad. Sci. U.S.A.*, **52**, 119–125 (1964).

Submarine pollination in seagrasses

THERE is great interest in water plants and their adaptation to the aquatic environment. Seagrasses, in the family Cymodoceaceae, are the only wholly marine group of flowering plants which carry out their entire life cycle in the sea[1]. The male and female flowers are borne on different plants. Consequently, their pollination mechanism presents intriguing problems, for terrestrial angiosperms, from which seagrasses presumably arose[2,3], shed dry pollen. This becomes hydrated only after alighting on the female stigma, where recognition events determine acceptance or rejection of the pollen[4]. These events seem to involve interactions between surface proteins or glycoproteins borne by both pollen and stigma[5]. The pollen proteins are released on to the stigma during normal pollination[6,7], but are lost from the surface whenever pollen is wetted[8]. The system found

in terrestrial species could thus hardly operate with submarine pollination. We have found that adaptations have occurred in both pollen and stigma of the seagrasses, accommodating the pollination system to the marine environment. Changes in the shape and form of the pollens, together with the loss of the outer wall layer, make it possible for the pollen to be carried in water currents as long, rope-like masses. The receptive stigma cells secrete a proteinaceous surface layer that is not dispersed in seawater, providing a suitable medium for trapping the pollen during submarine pollination.

We have used two species of sea nymph, *Amphibolis antarctica* and *A. griffithii*, found only on the temperate southern coast of Australia, and *Thalassodendron ciliatum*, widely distributed in tropical parts of the Indo-Pacific region, collected in Madagascar and Mauritius. No significant facts on the nature of their reproductive structures have been published since the pioneering studies of Ascherson[9] and Black[10], reviewed by den Hartog[3]. We have found in both genera that flowering occurs in early summer, and the plants are dioecious with separate male and female individuals. The flowers are hidden within the leaf sheaths. The paired anthers (Fig. 1a) extend upwards by peduncle elongation at maturity, when the entire structure is abscinded and floats free at the time of anther dehiscence. The female flowers remain hidden, the filamentous stigma branches elongate when receptive and protrude between the leaf sheaths (Fig. 2a).

The pollen grains are remarkable. They are elongate and confervoid, each more than 6 mm long, non-aperturate and at maturity they are coiled like a spring within the anther (Fig. 1b). They are trinucleate, with the nuclei located
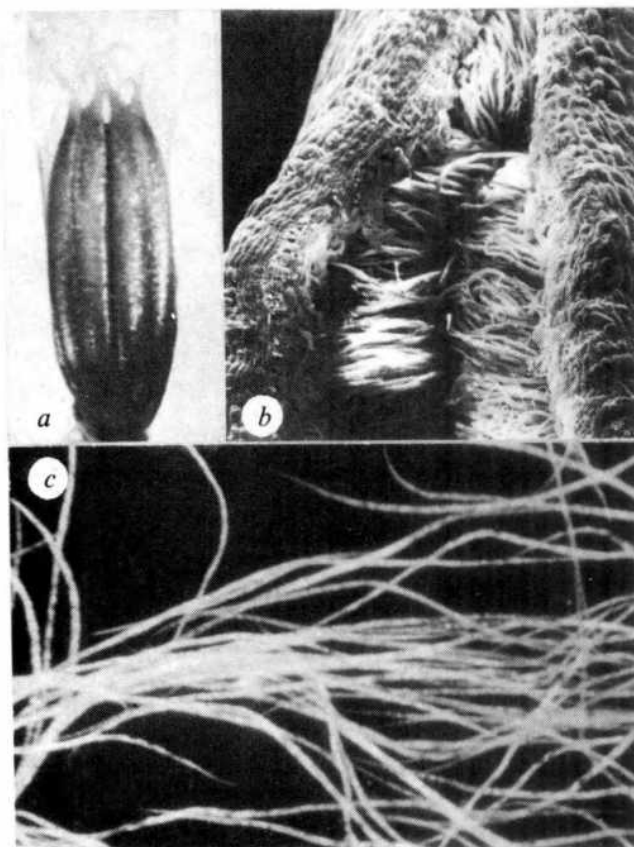
Fig. 1 Male flowers of the sea nymph, *Amphibolis antarctica*. *a*, Male inflorescence consisting of paired anthers fused to form a single dispersal unit (×4). *b*, Scanning electron micrograph of anther freeze dried at dehiscence, showing release of coiled pollen grains (×45). *c*, Vermiform pollen grains dehiscing from anther into seawater, revealed by the fluorochromatic reaction[19], a test for pollen viability. The cytoplasm of the grains shows intense green fluorescence (×80).

centrally. The walls are unusual in apparently having no outer exine, which in terrestrial plants is made of the polymer sporopollenin[11]. Cytochemical tests have revealed no trace of exine. In transverse sections, the walls of mature pollen stain a uniform red colour with toluidine blue, like the inner intine layer of terrestrial angiosperm pollens. Microspectrophotometric analysis of the red stain has revealed only the 565-nm peak characteristic of the intine, and not those of the green-stained exine (590 and 640 nm) of the terrestrial plants; fluorescence microscopy after staining with 0.01% auramine O at pH 5.0, which provides even higher resolution, has revealed absence of the yellow fluorescence characteristic of the exine (personal communication from J. Heslop-Harrison). Transmission electron microscopy of other aquatic plants[12] has demonstrated a progressive loss of exine, though in no case was it absent.

The grains adhere together, forming a rope-like structure when teased from a mature anther into seawater (Fig. 1c). The pollen is soft and flexuous, bending in water currents as it escapes from the anther. The surface appears smooth under scanning electron microscopy, and a faint reticulate pattern is evident at higher magnification. The tips of the grains grade into a pair of fine, curved hooks, which presumably enable the pollen grains to link together on dehiscence and to attach to the smooth stigma surface. As well as these adaptations to the marine environment, seagrass pollen has some features in common with land plants[13]. Seagrass pollen walls contain cytochemically-detectable acid phosphatase activity, except that the enzyme reaction product is distributed throughout the wall instead of being confined to the inner intine layer. Furthermore, esterase activity, usually restricted in terrestrial plants to the outer exine layer[14], occurs on the surface of the grains in immature anthers associated with the invasive tapetum but is absent in mature pollen.

In both genera, the paired female flowers each have a style bearing several stigma branches. The tips are coated, when receptive, with a surface layer which can be detected cytochemically by its esterase activity (Fig. 2b). Such layers are characteristic of both wet and dry stigmas of land plants, and have been implicated in pollen recognition[15,16]. In terrestrial species, the layer can be partially solubilised in buffer and detergent[5]. In *Amphibolis*, however, the esterase activity of the coating does not disperse readily in seawater. In stigma sections, enzyme reaction product is present in the outer walls of the epidermal cells which correspond to the papillae of terrestrial plants and in the extra-cuticular layer secreted from the cells on the surface (Fig. 2c).

Fertile fruits of *Amphibolis* are frequent, so that communication between the pollen released into the sea from male flowers and the stigma does indeed occur. Pollination takes place when the plants are totally submerged, and must involve unique adaptations for cell recognition. The surface of the flexuous pollen grains is adhesive, suggesting the presence of a coating, possibly glycoprotein in nature, that is either resistant to the solubilising action of seawater, or constantly being secreted from within the grains. Arabinogalactan proteins are secreted in large quantities into the medium from *Lolium multiflorum* endosperm cells in liquid culture[17], suggesting that such a mechanism is not without precedent.

Both *Amphibolis* and *Thalassodendron* are viviparous[3] and the single basal embryo develops directly into a plantlet. After fertilisation, special floral structures differentiate to enable the seedling to find a foothold on the sea bed. Vivipary of this kind is also found in other intertidal plants like mangroves and the monocotyledon *Cryptocoryne*[18].

We believe that the features we have described are those of a group of angiosperms that has adapted successfully to the marine environment by the loss or modification of terrestrial features.

S. C. DUCKER
R. B. KNOX

School of Botany,
University of Melbourne,
Parkville,
Victoria 3052,
Australia

Fig. 2 Female flowers of *Amphibolis antarctica*. *a*, Receptive stigma branches protrude from the leaf-like bracts (×2). *b*, Tips at higher magnification after cytochemical reaction for esterase, with α-naphthyl acetate as substrate in a simultaneous coupling reaction with hexazotised pararosanilin[7,15,16]. The red reaction product appears as a smooth coating on the extreme tips (×12). *c*, Stigma tip treated as (*b*), in transverse section, viewed by Nomarski interference contrast illumination, showing localisation of esterase in papillar (epidermal) cell walls and the outer extracuticular surface (pellicle). c, Cuticle. p, pellicle (×350).

1 Arber, A., *Water plants* (Cambridge University Press, 1920).
2 Sculthorpe, C. D., *The Biology of Aquatic Vascular Plants* (Arnold, London, 1967).
3 Hartog, C. den., *The Seagrasses of the World* (North-Holland, Amsterdam, 1970).
4 Heslop-Harrison, J., *A. Rev. Pl. Physiol.*, **26**, 403–425 (1975).
5 Knox, R. B., Clarke, A. E., Harrison, S., Smith, P., and Marchalonis, J. J., *Proc. natn. Acad. Sci. U.S.A.* (in the press).
6 Knox, R. B., and Heslop-Harrison, J., *J. Cell Sci.*, **9**, 239–251 (1971).
7 Heslop-Harrison, J., Knox, R. B., Heslop-Harrison, Y., and Mattsson, O., in *The Biology of the Male Gamete* (edit. by Duckett, J. G., and Racey, P. A.), 189–202 (Academic, London, 1975).
8 Stanley, R. G., and Linskens, H. F., *Physiol. Pl.*, **18**, 47–53 (1965).
9 Ascherson, P., *Linnaea*, **35**, 152–208 (1867–68).
10 Black J. M., *Trans. R. Soc. S. Aust.*, **37**, 1–5 (1913).
11 Brooks, J., and Shaw, G., in *Pollen Development and Physiology* (edit. by Heslop-Harrison, J.), 99–114 (Butterworth, London, 1972).
12 Pettitt, J. M., and Jermy, A. C., *Micron*, **5**, 377–405 (1975).
13 Knox, R. B., and Heslop-Harrison, J., *Nature*, **223**, 92–94 (1969).
14 Knox, R. B., Heslop-Harrison, J., and Heslop-Harrison, Y., in *The Biology of the Male Gamete* (edit. by Duckett, J. G., and Racey, P. A.), 177–187 (Academic, London, 1975).
15 Mattsson, O., Knox, R. B., Heslop-Harrison, J., and Heslop-Harrison, Y., *Nature*, **247**, 298–300 (1974).
16 Heslop-Harrison, J., Heslop-Harrison, Y., and Barber, J., *Proc. R. Soc.*, **B188**, 287–297 (1975).
17 Anderson, R., Clarke, A. E., Jermyn, M. A., Knox, R. B., and Stone, B. A., *Aust. J. Pl. Physiol.*, (in the press).
18 Goebel, K., *Organography of Plants* (Clarendon, Oxford, 1905).
19 Heslop-Harrison, J., and Heslop-Harrison, Y., *Stain Tech.*, **45**, 115–120 (1970).

matters arising

Nucleosynthesis and anomalous Xe and Kr in carbonaceous chondrites

BLACK[1] has suggested that the anomalous Xe isotopic composition known as carbonaceous chondrite fission (CCF) Xe results from a varient of the r-process and is not due to fission. To yield the CCF Xe isotopic composition, the r-process peak due to the neutron magic number 82 would need to occur at mass 136 rather than 130 as in standard Solar System material[2].

An actual r-process scenario has two difficulties. First, r-process calculations yield an abundance peak far too steep compared with the CCF data, when the abundance peak occurs close to the magic number on the valley of β stability (for the magic number N =82, β stability is at mass 138). Second, the time scale of an r-process resulting in a peak within two mass units of β stability is very long compared with that of an exploding object (the r-process time scale is determined by the β rates, which are relatively long near the valley of β stability); thus there is no plausible astrophysical site.

Another closely related possibility exists: a neutron capture process much faster than an s-process but operating under the same principle of competition between neutron capture and β decay. We have prepared a computer program to carry out such generalised neutron capture (n-process) studies[3]; the neutron capture cross sections required by the program code are derived from the Hauser–Feshbach code of Holmes et al[4]. This n-process may occur in novae, supernovae shocks or in explosive supernovae shells.

We have examined the predicted Xe and Kr relative (heavy) isotopic abundances as a function of neutron flux for neutron temperatures of 2–155 keV (Fig. 1). Also shown in Fig. 1 are the recent Allende experimental results of Anders et al[5,6] and the implied heavy Xe spectrum for Murchison[7] after a large fractionation correction (the nonfractionation corrected Murchison heavy Xe spectrum looks like the Allende heavy Xe spectrum). All other proposed heavy Xe spectra lie somewhere between these two curves. It is not possible to get a good fit for either the Anders et al. unfractionated Xe or the Kr data and, furthermore, the best

Fig. 1 The calculated isotopic abundances of Kr and Xe as a function of neutron flux. Also shown are Allende experimental results of Lewis et al[6] and the Murchison Xe results of Srinivasan et al[7] with their large fractionation correction. The calculations may fit the Murchison xenon results but they cannot fit the Lewis et al[6] Kr results, nor a heavy Xe spectrum with a large dip at 132. ●, Allende; ■, Murchison; neutron flux on an arbitrary scale: □, 1; ◇, 30; ○, 70; △, 300.

fit for each does not occur for the same flux. These results are for 155-keV neutrons, similar results obtain at other neutron energies down to 2 keV. Near β stability the cross section computer code is good to a factor of order two (S. E. Woosley, personal communication); greater changes in neutron cross sections would be needed and specifically selected to get a good fit to the unfractionated Allende results. Such a selected set of cross section changes is not likely. Note, however, that the heavy Xe spectrum corrected for a large implied fractionation is not very different from a possible n-process. If further experimental work indeed shows the carbonaceous chondrite heavy Xe spectrum to be relatively smooth without a dip at 132, then perhaps the results can be explained with some sort of an n-process. Unfortunately, the Kr results cannot be fitted, but they are probably less reliable than the Xe results (E. Anders, personal communication). If the heavy Kr spectrum of Anders et al. is verified then a steady-state n-process explanation is no longer possible.

There is another important point with regards to the Allende xenon results of Anders et al[5] and Lewis et al[6]. Pepin[8] has pointed out that enriched light Xe isotopes seem to be a separate component from the heavy isotopes. From examination of the relevant nuclear processes, it seems clear that the light Xe anomalies cannot be fitted with spallation reactions —such reactions always produce $^{126}Xe > ^{124}Xe$ for any plausible seed distribution. Perhaps the explanation of the anomalous light isotopes lies in either the p-process[9] or in fractionation[10]. If the light xenon component is due to the p-process then ^{130}Xe might be effected, which would change the normalisation[11] on the heavy Xe, tending to decrease the dip at 132. Thus both the p-process and fractionation, if consistently applied, will tend to smooth the heavy Xe.

These n-process computations were done using the long time steady-state solution and thus are independent of initial composition. We conclude that a classical r-process cannot yield the observed Xe–Kr isotopic distribution. A long time n-process would fit only if there was no dip in the Xe spectrum at 132 and if the carbonaceous chondrite Kr spectrum were very different from that observed by Anders et al. It is also possible however, that the anomalous gas distributions could result from the addition of only a few or several neutrons to already existing heavy seed nuclei. (One can get a good fit if the number of free parameters equals the number of isotopic abundances to be fitted; such a procedure is not very useful.)

We are currently examining the results of adding a few neutrons to preexisting seed and limiting the free parameters by using plausible theoretical

pre-supernovae isotopic compositions and by investigating what other observable effects would result from such a neutron burst. These may well be limited to the other noble gases; effects on other elements may be masked by their much greater normal abundance. The superheavy element suggestion of Anders *et al.*[5] of course remains a possibility.

We thank D. Black and E. Anders for useful comments. This work was supported in part by NASA and the NSF.

J. B. BLAKE

Space Sciences Laboratory,
The Aerospace Corporation,
Los Angeles, California 90009

D. N. SCHRAMM

Enrico Fermi Institute,
University of Chicago,
Chicago, Illinois 60637

1 Black, D. C. *Nature*, 253, 417–419 (1975).
2 Seeger, P. A., Fowler, W. A., and Clayton, D. D., *Astrophys J. Suppl.*, 11, 121–166 (1965).
3 Blake, J. B., and Schramm, D. N., *Astrophys J.* (in the press).
4 Holmes, J., Woosley, S. E., Fowler, W. A., and Zimmerman, B. A., *Nucl. Atomic Data Tables* (in the press).
5 Anders, E., Higuchi, H., Gws, J., Takahashi, H., and Morgan, J. W., *Science*, 190, 1262–1271 (1975).
6 Lewis, R. S., Srinivasan, B., and Anders, E., *Science*, 190, 1251–1262 (1975).
7 Srinivasan, B., Gros, T., and Anders, E., *J. geophys. Res.* (in the press).
8 Pepin, R. O., *EOS, Trans. Am. geophys. Un.*, 57, 278 (1976).
9 Clayton, D. D., *Geochim. cosmochim. Acta*, 40, 563–565, (1976).
10 Sabu, D. D., and Manuel, O.K., *EOS, Trans. Am. geophys. Un.*, 57, 278 (1976); *Nature*, 262, 28–32 (1976).

Elevation of selenium levels in air by xerography

THE report on this subject which appeared in *Nature*[1] raises questions of methodology and relevance which I discuss here.

Although the authors calculate airborne selenium concentrations between 2×10^{-8} and 6×10^{-8} g m^{-3}, a simple calculation based on the data presented yields a concentration range 4.9×10^{-9}–14.7×10^{-9} g m^{-3}. Of more concern is the sensitivity of the analytical methods used. Their prime reference[2] reports "precision and accuracy are good for Se levels down to about 0.1 p.p.m. (1×10^{-7} g per determination) and acceptable for many purposes to 0.02 p.p.m. (2×10^{-8} g)" and reagent blanks ranging from 2 to 2.3×10^{-8} g selenium per determination. Using Olson's method[2], Harkin *et al.*[1] reported collection of from 5×10^{-9} to 15×10^{-8} g of airborne selenium in a xerography room. These values are clearly below Olson's own sensitivity limit and are even below Olson's blanks[2]. Harkin has made assurances (personal communication) that Olson's methods have been improved on and that the values reported are in excess of blanks, but the precision and

accuracy of Harkin's data cannot be verified by the methods cited.

Even assuming Harkin's results are valid, some perspective is appropriate. The highest selenium concentration reported[1] (6×10^{-8} g m^{-3}) is 3,333 times lower than the US Occupational Safety and Health Agency (OSHA) limit[3] (2×10^{-4} g m^{-3}). If one assumes a human minute volume of 8.2 l (ref. 4) and total absorption of inspired selenium[5] at the highest level reported, the total selenium ingested by inhalation in an 8-h day (2.4×10^{-7} g) would be 27 times less than that ingested by consuming a single 23-g slice of white bread (6.4×10^{-6} g)[6]. Selenium, incidentally, is considered as a nutritionally essential trace element[7].

We have measured selenium emissions from Xerox machines and have found them to be several orders of magnitude below the OSHA limit and below the usable limit of our analytical methods. Since the question has been raised, however, we have initiated a renewed effort externally to ensure that we have the most accurate data possible.

R. A. PARENT

Product Safety,
Xerox Corporation,
Rochester, New York 14644

1 Harkin, J. M., Dong, A., and Chesters, G., *Nature*, 259, 204 (1976).
2 Olson, O. E., *J. Ass. off. agric. Chem.*, 52, 627–634 (1969).
3 *Federal Register*, 39 (125), 23542, Washington, DC. June 27, 1974.
4 Hurtado, A. in *Handbook of Physiology* Sect. IV (edit. by Dill D. B.), 855 (Waverley, Baltimore, 1964).
5 Dutkiewicz, T., Balcerska, J., and Dutkiewicz, B., *Bromatol. Chem. Toksykol.* 4 (2), 185–191 (1971); *Chem. Abstr.*, 76, 10843k (1972).
6 Scott, M. L. in *Organic Selenium Compounds: their Chemistry and Biology* (edit. by Klayman, D. L., and Günther, W. H. H.), 652 (Wiley-Interscience, New York, 1973).
7 Frost, D. V., and Lish, P. M., *A. Rev. Pharmac.*, 15, 259–284 (1975).

HARKIN ET AL. REPLY—Fluorimetric analysis of selenium using 2,3-diaminonaphthalene (DAN) is easily sensitive enough to determine the levels we collected from air. The sensitivity claimed[1] for the method is very conservative and can be improved by slight modifications of the basic procedure[2]. These include recrystallisation of the DAN reagent, extraction of the fluorescing Se–DAN complex with cyclohexane rather than with decahydronaphthalene, and extension of the digestion time with perchloric–nitric acid to 90 min following the first appearance of perchloric acid fumes. Sensitivities of 1 ng and <4 ng have been claimed for blank analyses by others[3,4]. Our blank values lay reproducibly at 4.7 ± 0.7 ng Se.

The question[5] of the sensitivity of the method is not really pertinent, since the figures we reported[6] were for levels above values measured in control experiments using air samples collected in laboratories in the same building as our copying room. Because the values were low for single-day air samples, to

increase the reliability of the analyses, cumulative samples taken over 4 d were actually analysed, but values were supposed to be expressed on a daily basis. Omission of the word "daily" after the values of 0.005–0.015 ng in the report[6] made them seem to be the 4-d total rather than single-day averages. The value of 20–60 ng Se per m^3 of air is the correct, rounded-off figure for the airborne Se levels.

Tolerable limits for exposure to selenium depends on the Se species concerned, and some reappraisal of current values may be warranted. The time-weighted average concentration of 0.2 mg m^{-3} permitted by the OSHA for elemental selenium and selenium oxides in air differs from the value for hydrogen selenide, which at 0.05 p.p.m. is 200 times lower than that for hydrogen cyanide, which is generally considered to be the most highly toxic common airborne inorganic compound[7]. Russian workers recently suggested that the permissible limits for SeO$_2$ should be reduced to 0.1 μg m^{-3} in air for a single exposure and to 0.05 μg m^{-3} in air for average daily exposure[8].

Comparison of inhaled and ingested selenium is of doubtful validity in the absence of toxicological work directed at this specific question. There is, in fact, evidence that selenium administered in the oxidised form can be excreted in part by respiration as volatile methylated derivatives[9].

Although excess selenium may be harmful under ordinary circumstances, its beneficial effects in counteracting heavy metal toxicity should not be overlooked[10–13]. Since there are many questions of selenium biochemistry and metabolism still unanswered, the effects of minute volatile emissions from copiers cannot be predicted with confidence. We appreciate the efforts of copier manufacturers to ensure that these emissions will be kept to a minimum and checked by independent analyses.

University of Wisconsin

1 Olson, O. E., *J. Ass. off. agric. Chem.*, 52, 627–634 (1969).
2 Olson, O. E., Palmer, I. S., and Cary, E. E., *J. Ass. off. agric. Chem.*, 58, 117–121 (1975).
3 Hoffman, I., Westerby, R. J., and Hidiroglou, M., *J. Ass. off. agric. Chem.*, 51, 1039–1042 (1968).
4 Levesque, M., and Vendette, E. D., *Can. J. Soil Sci.*, 51, 85–93 (1971).
5 Parent, R. A., *Nature*, 263, 708 (1976).
6 Harkin, J. M., Dong, A., and Chesters, G., *Nature*, 259, 204 (1976).
7 Christensen, H. E., Luginbyhl, T. T., and Carroll, B. S., *Registry of Toxic Effects of Chemical Substances* (US Dept of Health, Education and Welfare, Washington, DC. 1975).
8 Selyankina, K. P., Yakhimovich, N. P., Lenchenko, V. G., Alekseeva, L. S., and Pochashev, E. N., *Gig. Sanit.*, 10, 14–18 (1975).
9 Ganther, H. E., Levander, O. A., and Baumann, C. A., *J. Nut.*, 88, 55–60 (1966).
10 Stoewsand, T. S., Bache, C. A., and Lisk, D. J., *Bull. environ. Contam. Tox.*, 11, 152–156 (1974).
11 Ganther, H. E., *et al.*, *Science*, 175, 1122–1124 (1972).
12 Ganther, H. E., Wagner, P. A., Sunde, M.L., and Hoekstra, W. G., in *Trace Substances in Environmental Health*, VI (edit. by Hemphill, D. D.), 247–252 (University of Missouri Press, Columbia, 1973).
13 Ganther, H. E., and Sunde, M. L., *J. Food Sci.*, 39, 1–5 (1974).

reviews

Delights of opisthobranch molluscs

<div style="text-align: right">V. Fretter</div>

Biology of Opisthobranch Molluscs.
Vol. 1. By T. E. Thompson. Pp. 207.
(The Ray Society, c/o British Museum
(Natural History): London, 1976.)
£15.

OPISTHOBRANCHS, the unexpected prog-
eny of shelled gastropods, have relin-
quished their protective encasement to
reveal a grace of form and colour un-
surpassed by other molluscs. Yet most
have remained relatively unknown to
the malacologist because of their re-
tiring and migratory habits and their
fondness for the immediate sublittoral,
an area difficult to explore before the
days of SCUBA diving. This is the
first book related to British opistho-
branchs for over a hundred years. It
is in two sections: the first, dealing
with the general biology of the group,
includes some magnificent colour
photographs and drawings of external
features of worldwide forms; the
second is a systematic treatment of
British species, excluding nudibranchs
(which are to be dealt with in a second
volume), the parasitic pyramidellids
and the planktonic pteropods.

No comprehensive account of the
biological activities of sea slugs has
been written previously and in the first
section the author successfully con-
veys his interests in, and personal ex-
perience of the living animals. The
text is lucid and not loaded with ex-
cessive detail so that its contents are
easily accessible to the non-specialist.
It comprises a series of chapters deal-
ing in a general way with various as-
pects of opisthobranch biology from
the point of view of functional ana-
tomy; these constitute one of the out-
standing features of the book.

These aspects are dealt with almost
on the 'type' system since each con-
sists largely of summaries of the
author's own research on a number of
species. Without wishing to detract in
any way from the value of this, I wish
that the relationship to other opistho-
branch research had been explained
more fully to provide a better under-
standing of the biology of the group
as a whole, although, as opisthobranchs
have radiated so divergently, this may
be too difficult a task. I wish, too, that
the coverage of the systems had been
complete—reluctance to enter the field

of neurophysiology, where opistho-
branchs have furnished so much im-
portant material, is easily understood,
but there are other aspects of the
nervous system than the physiological;
and there are other systems in regard
to which the views of so experienced
and distinguished a worker would have
been welcomed.

The systematic section summarises
characters of all orders, provides keys
to British species dealt with in this
volume and a description of each, in-
cluding comments on the living animal.

This will be an invaluable guide to the
taxonomist. He, like others, would be
helped by a subject index.

Illustrations of anatomy are from
the author's published works, those of
co-workers and redrawings of the
anatomy of even some of the most
accessible species from classical illus-
trations published as far back as 1845.

---

*Dr Fretter is a former Reader in the
Department of Zoology at the Univer-
sity of Reading, UK.*

Meditations on a cell cycle

*Multiplication and Division in Mam-
malian Cells.* By Renato Baserga. Pp.
xii+239. (Dekker: New York and
Basel, May 1976.) SFr. 78.

PROFESSOR BASERGA'S book combines
those rare qualities of being readable,
informative and very often amusing.
The book makes no claims to be an
exhaustive study of the multiplication
of mammalian cells; ample references
direct the reader to review articles of
greater depth.

A major part of the monograph is
naturally taken up with a review of the
cell cycle and aficionados of the
relevant literature will recognise
Baserga's rather personal views on cell
cycle controls. Baserga marshalls a
number of arguments (none of them
convincing) to support the widely held
idea that G_0 cells are not merely cells
with a long G_1 which may be shor-
tened by suitable stimulae, but are a
separate class of cells awaiting their
fairy godmother to trigger them back
into the cycle. This is not just a seman-
tic argument as it determines whether
mitogenic stimulae are regarded as
triggers initiating new processes or ac-
celerators of ongoing events. I see
nothing wrong with the view that G_1
may be infinitely expandable to include
G_0. Unfortunately the cell cycle is often
drawn like the face of a clock with the
cell cycle phases apparently given defi-
nite durations. This ignores the fact
that variation in the length of cell cycle
phases (especially G_1) is the rule rather
than the exception, and failure to re-

cognise this has placed enormous
psychological constraints on the under-
standing of cell proliferation. The
criteria Baserga cites for distinguishing
between cells in G_1 and cells in G_0
result from sampling cells from widely
separate parts of a continuum of pro-
liferation rates.

The other hobby-horse Baserga rides
round the cell cycle is the idea that
new genes must be activated to trigger
cells from quiescence to rapid growth.
This may be true but I find his evi-
dence involving changes in the physical
structure of chromatin, chromatin tem-
plate activity and nuclear acidic pro-
teins far from conclusive. Recent
mRNA hybridisation studies show that
at most only 5% of the approximately
10^4 separate mRNA sequences differ
in resting and growing cells. As only
5–10% of the mammalian genome is
transcribed this difference represents
less than 0.5% of the total genome. I
find it difficult to believe that the large
scale changes in chromatin with which
he makes such a great play) are all
devoted to the derepression of such a
minute fraction of the genome.

This being said, much of the book
is good sense; anyone involved in cell
proliferation could do well to read it
if only as an antidote to the turgid
prose and thought behind much of the
burgeoning hard core of the cell pro-
liferation literature. **Robert Shields**

---

*Robert Shields is a Research Fellow
at the Imperial Cancer Research Fund,
London, UK.*

•Monoamine oxidase

Monoamine Oxidase and Its Inhibition.
(Ciba Foundation Symposium 39 (new
series).) (In Honour of Mary L. C.
Bernheim.) Pp. xii+415. (Elsevier/
Excerpta Medica/North Holland;
Amsterdam, Oxford and New York,
1976.) Dfl.78; $29.95.

THE present resurgence of interest in
monoamine oxidase (MAO) and its in-
hibitors (MAOIs) stems directly from
the discovery that at least two func-
tionally distinct forms of the enzyme
exist in many mammalian tissues, in-
cluding the human brain. The prob-
lem is now to assess the physiological
and pharmacological significance of
this finding particularly with regard to
the use of MAOIs in the treatment of
certain affective disorders. This volume
is intended to update and assemble
some of the more recent findings in the
field.

Almost without exception, the papers
comprising this book are of a high
standard. The first five articles deal
with the nature and properties of the
multiple enzyme forms together with
the mechanism of action of acetylenic
MAOIs. The second section catalogues
a series of re-investigations into the
physiological significance of MAO
using MAOIs that more or less speci-
fically inhibit one or other of the en-
zyme forms. A final group of seven
papers deals with the evaluation of
these newer MAOIs in the treatment of
certain types of depressive illness and
with also the possible usefulness of
changes in human blood platelet MAO
activity as indicators of such disease
states as depression, migraine and
schizophrenia.

The most interesting and revealing
portions of this volume are the 125
pages of discussion. A feeling of be-
wilderment and uncertainty is con-
stantly surfacing throughout these
pages. Is the 'amine hypothesis' of
depressive illness the correct model?
Against all preconceptions, may anti-
depressant drugs act by reducing
and not increasing the effective con-
centration of pharmacologically active
amines at the receptor site? Perhaps
the site of action of MAOIs is not in
the brain but in the periphery: the
accumulated amines may then pene-
trate the blood–brain barrier. Are
MAOIs effective antidepressants be-
cause they inhibit MAO or is it be-
cause they inhibit amine synthesis by a
feedback mechanism or perhaps block
the uptake of amines by cellular
elements in the brain? Is the ability of
the tricyclic antidepressant drugs to
inhibit one of the two forms of MAO
with a K_i of about 10^{-5} M a factor in
their mood-elevating properties? Is it
even possible to make meaningful
studies on the role of MAO *in vivo*
when that fraction associated with
nerve endings is less than 5% of the
total? These and many other problems
are discussed in this fascinating volume.

It seems absurd that so little is
known about the role of an enzyme
originally characterised as long ago as
1928 by Mary Bernheim. There is one
inescapable conclusion that may be
drawn from this book: until the aetio-
logy of depressive illness is understood
and the biochemists and clinical psy-
chiatrists achieve some standardisation
in their respective disciplines, the treat-
ment of affective disorders with MAOIs
will continue to be a 'hit-and-miss'
affair.   **Godfrey G. S. Collins**

*Godfrey Collins is a Lecturer in the
Department of Pharmacology at the
School of Pharmacy, University of
London, UK.*

Mammalian gluconeogenesis

*Gluconeogenesis: Its Regulation in
Mammalian Species.* Edited by Richard
W. Hanson and Myron A. Mehlman.
Pp. xxvi+592. (Wiley-Interscience:
New York and London, 1976) $36.40;
£18.60.

THIS volume of 15 essays is dedicated
to Professor Henry A. Lardy for his
outstanding contributions to bio-
chemistry–impressively listed in the
introduction. The editors have sought
to compile the most significant infor-
mation gleaned from 25 years of re-
search on gluconeogenesis, an area in
which Professor Lardy has played an
important part. The chapters are group-
ed into four sections: the enzymology
of gluconeogenesis; the involvement of
the mitochondrion; the experimental
approach to gluconeogenesis in liver
and kidney; and gluconeogenesis in
man. The sections seem to cover the
main areas of current interest: the
properties and regulation of the en-
zymes unique to glucose production
that do not function in glycolysis; the
vital question of the concentration of
substrates actually available to these
enzymes; the significance of the
intracellular location of the enzymes,
particularly phosphoenolpyruvate car-
boxykinase (PEP-CK); transmitochon-
drial membrane carriage of reducing
equivalents; the interaction and com-
petition for energy and substrate
between urea synthesis and gluconeo-
genesis; the different manner of control
of gluconeogenesis in liver and kidney;
and last but by no means least what it
all means for *homo sapiens.*

As the editors point out we now know
a great deal about the pathway of glu-
coneogenesis from various substrates,
but considerably less about how the
process is controlled. It remains un-
clear whether for glucose formation
from lactate and alanine pyruvate
carboxylase or PEP-CK constitutes the
pacesetting enzyme. The activity of the
former is strongly regulated by acetyl-
CoA, but it is pointed out that we can
at present only guess at the concentra-
tions of CoA esters available to the
enzyme *in vivo*. PEP-CK exists in at
least two forms and possesses a varied
distribution between mitochondrion and
cytoplasm in different species. The level
of the cytoplasmic form in particular
responds to a number of hormones.
Although the varied distribution is
described by several authors and the
range of activities in different species
compared, there seems no suggestion of
what the rationale for this diversity
may be. The enigma of an apparent
K_m for PEP-CK towards oxaloacetate
far in excess of available concentra-
tions, resolved for the cytoplasmic
enzyme by Ballard (*Biochem. J.* **120,**
809, 1970), remains an acute problem
for the mitochondrial enzyme since the
overall mitochondrial concentration of
oxaloacetate seems so low. It was Lardy
himself who pointed out the virtue of
a cytoplasmic location of PEP-CK if
it led to oxaloacetate transfer from the
mitochondrion as malate, since this
simultaneously resulted in 2H transfer
to the cytoplasm needed at the stage
of triose phosphate formation when
gluconeogenesis is proceeding from
substrate not initially undergoing oxi-
dation. It is interesting to learn that in
pigeon liver, where phosphoenolpyru-
vate synthesis is almost entirely mito-
chondrial, negligible glucose synthesis
from either pyruvate or alanine occurs.
Except in relation to man, little con-
sideration seems to be given to the role
of glycerol as a gluconeogenic sub-
strate, yet in prolonged fasting it be-
comes quantitatively important and
could also be a source of reducing
equivalents for other precursors.

This volume will be indispensable to
all workers in the field and is to be
recommended to anyone interested in
gluconeogenesis. It is a pity that a
number of the chapters have no sum-
mary section, but, perhaps to compen-
sate, the editors have provided their
own synopsis of the contributors'
main points and the current state of
the art.   **K. L. Manchester**

*K. L. Manchester is Professor of Bio-
chemistry at the University of the
Witwatersrand, Johannesburg, South
Africa.*

obituary

Hermann Träuble was born on April 7, 1932 at Nellingen, near Ulm, and died on July 3, 1976.

It was about nine years ago that Hermann Träuble first put his head round my door with a rhetorical "Am I disturbing you?" This pleasant young man who stood there before me, with a modest and yet forceful presence, was by no means an unknown quantity to me. His teacher and supervisor at the Max-Planck-Institut in Stuttgart, Alfred Seeger, had notified me of his visit, with the words which Mozart is said to have used of Beethoven: "Keep an eye on him, one of these days he'll make the news".

Hermann Träuble had already made the news. Together with Uwe Essmann he had successfully carried out an experiment which others had only dreamed of—the direct, visible demonstration of a quantum effect. This experiment is so typical of its performer's straightforward, goal-directed methods of thought and work that I would like to outline it here in a few words. I shall try to relate the story in the way which Hermann Träuble himself described it to me.

The experiment was an attempt to demonstrate the quantisation of magnetic lines of force in a particular class of superconductors, the so-called Class II superconductors, which are characterised by a particular kind of behaviour in a strong magnetic field.

How does one make the paths of magnetic flux lines visible? The answer is known even to schoolboys: use fine iron filings—but just how fine must the filings be, to demonstrate the quantum phenomenon? One would answer intuitively: "As fine as possible, best of all a dust made up of individual atoms or small aggregates of atoms." But that would be a hasty answer, for we must recall that ferromagnetism is not a property of individual atoms, and it can only appear in ordered crystalline structures.

The question should therefore more precisely run: how small must the particles be, to respond to a quantum effect, and how large must they be to act ferromagnetically and to appear on the screen of an electron microscope.

The two young scientists grasped the bull by the horns, and simply started to experiment. Using an inert-gas atmosphere, they succeeded in producing tiny crystals with an edge length of ~ 0.000001", and then in condensing these onto the surface of a semiconductor in a magnetic field. The circulating currents in the semiconductor induced by the magnetic field are divided up into tiny vortices, which rotate about an axis parallel to the magnetic lines of flux. The minute ferromagnets deposited themselves preferentially onto this lattice of 'quanta of magnetic flux' and build up crystal nuclei, which could be seen under an electron microscope.

Result: a direct, macroscopic observation of a microscopic quantum effect!

The Physics Prize of the Göttingen Academy of Sciences awarded to these two scientists was a clear expression of the recognition with which the world of learning received this great achievement. For Hermann Träuble this was, however, not the first distinction. A short while previously he had received the Masing Preis of the German Metallurgical Society for his work on the magnetisation and hysteresis in ferromagnetic single crystals, which he had carried out as part of his doctorate.

Now he was standing in front of me. All he wanted was a post as an assistant —a man who had a brilliant career in solid-state physics in front of him and to whom a university chair was already open. All he wanted was a small laboratory, he assured me, maybe a hundred square feet—after taking a critical look at the dimensions of my study. He had taken it into his head to get into biophysics and a true Swabe can never be deflected from his intended course.

Hermann Träuble's interests in Göttingen were particularly directed at the structure and function of lipid membranes. In 1971 he published a comprehensive article in *Die Naturwissenschaften*: 'Phasenumwandlungen in Lipiden', whose subtitle 'Possible Switching Processes in Biological Membranes' already looked to the future. He hitched his own work onto that of Dennis Chapman, who had demonstrated calorimetrically the phase transitions between crystalline and liquidcrystalline states. Hermann pioneered the development of new and above all fluorimetric methods which made it possible to follow these rapid co-operative transformations. Thus he was able to investigate not only the static but also the dynamic aspects of these processes. The high speed of the phase transformations is closely connected with their potential significance as biological switching processes. Collaboration with Hansjörg Eibl made possible a comprehensive study of those membranes whose lipid constituents were already well known. Models were developed, their electrostatic interactions were studied, and the findings compared in each instance with similar or related effects in biological systems.

These were the pioneer years of a new 'membrane biophysics', full of excitement and creativity. Seldom did Hermann go home before daybreak, and even then he was always back in the Institute at five to two in the afternoon—just in time to grab something to eat in the canteen.

Wherever co-operation was offered, it was taken up. Peter Overath contributed his wide experience of the Coli membrane, and through this collaboration a detailed picture of the structure and phase transformations of bacterial membranes was put together, backed up by studies with models. Together with Erich Sackmann, who was familiar with the technique of electron spin resonance, he attacked the problem of lateral diffusion of the lipid molecules in the membrane. This turned out to be a neck-and-neck race with the group of Harden McConnell in Stanford.

He also applied his research to fundamental biological problems— a classic example of this is the paper with Hansjörg Eibl and Hideo Sawada "Respiration—a Critical Phenomenon?"; it is still too early to see the significance of his investigations for the final understanding of nerve membranes. Not without good reason was Hermann Träuble a much sought-after discussion partner in numerous workshops in the Neurosciences Research Program in Boston, USA. Listening to his lectures was a pure delight. His vital asset was that, where others had only answers, he always posed the right question.

The walls of his original laboratory had long since become too narrow. In 1974 the Max-Planck-Gesellschaft received him as a Scientific Member, and only a few weeks before his death he was appointed to a directorship in the Institute for Biophysical Chemistry in Göttingen. His research group had grown uninterruptedly in recent years, and includes today a dozen scientists, who will undoubtedly go on to extend the ideas which they and Träuble developed.

Neither was there any lack of new honours. The Faculty of Human Medicine of the University of Giessen

• awarded Hermann Träuble the Ludwig Schunk Prize in 1972 and the German Bunsen Society for Physical Chemistry the Bodenstein Prize in 1975.

The man, just 44, was called away at the zenith of his career. We shall not meet Herman Träuble again in the corridors of the Institute, or be enlivened at our tea colloquia by his burning enthusiasm and swabish charme; we shall not race together down the ski slopes in January, and I shall never have another chance of digging with him for fossils in California. Certainly, death is an integral part of life. But this insight cannot take the pain away.

Manfred Eigen

Ian Macpherson, head of the Department of Virology at the Imperial Cancer Research Fund Laboratories, London, died on September 11, 1976 at the age of 46, in the middle of a distinguished career. He was married with three children.

Macpherson was born and educated in Scotland, as a Foundation Scholar at George Herriots, and after at Edinburgh University, leading to a Carnegie Fellowship and a Ph.D. in virology in 1955. He then worked at the Microbiological Research Establishment, Porton, for four years before moving to Glasgow in 1959. After a year in the Department of Genetics, he became a founder member of the new Institute of Virology and of the Medical Research Council's Experimental Virus Research Unit. Except for a year as an Eleanor Roosevelt Cancer Fellow at the Wistar Institute, Philadelphia, he remained in Glasgow until 1968 when he was appointed to the senior staff of the Imperial Cancer Research Fund Laboratories in London to head a substantial department.

Macpherson began to study tumour viruses in Glasgow and it was for his unique and pioneering contributions to the study of virus transformed cells that he will be remembered. Of great importance was the isolation of the first line of cultured cells (BHK 21 cells) which could be used for precise quantitative evaluation of the transformation process, and which allowed direct comparison of transformed and untransformed cells from the same clonal population. Then came the discovery, with Montagnier, of the agar suspension assay system, which, with its various modifications, is of profound importance in tumour cell biology. Macpherson also isolated the first revertants of transformed cells, thus paving the way for the use by others of this important method for studying the transformed cell phenotype. In recent years Macpherson and his colleagues in London concentrated on the altered surface chemistry of transformed cells, and in particular developed the use of conditional mutants for elucidating the role of the virus.

Through Macpherson's wide reputation, his laboratory attracted visitors from far and near. He lead his group from the laboratory bench, not only by his originality and knowledge but by his experimental ability. Those associated with him soon discovered another aspect of his character as well—his impish sense of humour. This was often expressed in remarkable collages in which he delighted to ridicule any pompous and unjustified aggrandisment.

Ian Macpherson's hobby was golf, in which he also excelled, and it was on the golf course that he died.

The eminent organic chemist **George O. Curme, Jr** died on July 28 at his summer home at Oaks Bluff, Massachusetts. He was born in Mount Vernon, Iowa in 1888, the son of a professor of German who was also a renowned grammarian. Curme graduated at Northwestern University, then as a graduate student moved from Harvard to the University of Chicago, where he obtained his Ph.D. Next he went to Germany, then the Mecca of organic chemists (it has been said that in the first half of the century all organic laboratories were built pointing towards Berlin) to study under Fritz Haber and Emil Fischer at the Kaiser Wilhelm Institute and the University of Berlin.

On returning to America he worked from 1914–1920 as a research fellow at the Mellon Institute, and it was here that he started on the work which founded the synthetic aliphatic organic chemistry industry in America. (Aliphatic chemistry deals with chain-like compounds of carbon, as opposed to aromatic chemistry which deals with ring compounds.) He initially sought a method for the commercial manufacture of acetylene; he later widened his horizons to many other aliphatic compounds.

In 1917 he started to work for Union Carbide, with whom he was to remain (eventually becoming a director) until his retirement. From the Union Carbide laboratories in 1925 he announced the discovery of a process to make ethylene glycol, which soon replaced alcohol for use as an antifreeze in cars' radiators, and which rapidly developed into a multi-million dollar industry. Dr Curme continued his research, producing a new synthesis of ethyl alcohol, and later playing an important part in the wartime synthetic rubber programme.

Among the many awards he was given was the Willard Gibbs Medal, bestowed on him by the American Chemical Society in 1944 for his fundamental role in bringing "leadership in organic chemistry from Germany to the U.S.A."

---

announcements

Meetings

November 17–19, **ACTH and Related Peptides: Structure, Regulation and Action,** New York (Mr Charles Roarty, The Barbizon-Plaza Hotel, 106 Central Park South, New York, New York 10019).

Winter Gordon Research Conferences: January 3–7, **Deformation and Failure Mechanics in Polymer Composites; Bacterial Cell Surfaces.**
January 10–14, **Polymers; Organic Thin Films and Solid Surfaces.**

January 17–21, **Electrochemistry; Agricultural Science.**
January 24–28, **Chemical Oceanography.**
January 31–February 4, **Hormone Action.**
(Alexander M. Cruickshank, Director Gordon Research Conferences, Pastore Chemical Laboratory, University of Rhode Island, Kingston, Rhode Island 02881).

June 1–3, **Frequency Control,** Atlantic City, New Jersey (Deadline for abstracts: January 21) (Commander, US Army Electronics Command, ATTN: DRSEL-TL-MF (Dr J. R. Vig), Fort Monmouth, New Jersey 07703).

September 5–9, **Precise Electrical Measurements,** Brighton, Sussex (Conference Secretary, The IEE, Savoy Place, London WC2R 0BL).

September 5–10, **Geochronology, Cosmochronology and Isotope Geology,** Pisa, Italy (Deadline for abstracts: May 15) (Gabriella Bonadonna, C.N.R.-Laboratorio di Geochronologia, Via Cardinale Maffi, 36, 56100 Pisa, Italy).

nature

October 28, 1976

Dear Mrs Williams... HELP

BRITAIN'S science funding through the research councils has got itself into a real mess these past few weeks. Shirley Williams, Secretary of State for Education and Science, whose stock is already high amongst both administrators and scientists, could well earn herself even more credit by managing to sort out a logjam which no one else has the political clout to do.

Faced like every other body with the need for economic restraint, the Advisory Board for the Research Councils (ABRC) has been saying for some time now, and most notably in its Second Report published about six months ago, that if expenditure is allowed to grow at all it is the research councils other than the Science Research Council (SRC) which will get any growth that is going. Indeed, to help fund this the SRC has had to expect to spend 2% less per year over the years to 1980. This makes the period 1973–80 one in which the spending power of the SRC in real terms is expected to fall by one-sixth.

Within that overall decline, moreover, there were to be favoured areas. Engineering research and associated schemes such as the Teaching Company, marine technology and polymer engineering, were to grow by more than 50% in the years 1976–81. Science Board research comprising biology, chemistry, mathematics and physics outside astronomy and nuclear physics is allowed a 10% growth in the same period, although all of this and more will be taken up by expansion of central facilities. On the other hand nuclear physics expenditure (at present £41 million per annum, half domestic, half at CERN) will drop to £29.5 million, with all domestic expenditure nearly halved. And expenditure in astronomy and space research will drop from £26 million to £20 million, again with the subscription to the European Space Agency assuming an increased importance. Whatever abuse may be hurled at SRC, it has certainly shown itself capable of rapid response to outside pressures for changes in direction—some of these switches mean that laboratories are having to change their rôle entirely.

Now all of this makes for a most remarkable contrast in reading between the SRC's annual report, just published (HMSO: £1.75) and the reports of the Natural Environment Research Council — NERC (HMSO; £2.65) — and the Medical Research Council — MRC (HMSO: £2.75). (Even the price difference tells the story). Whereas NERC and MRC, although generally worried by economic stringency, seem to find their biggest problems in such matters as adapting to the customer-contractor relationship and giving their researchers a stable career, SRC looks burdened down on almost all sides. But it is what has happened since the end of the reporting year (31st March, 1976) which now has everyone so much on edge. Changes in the value of the pound have put up the cost of the international subscriptions to just those agencies such as CERN and ESA that are meant to take the load off the domestic budget. Of the £6 million more that SRC needs to find, the Treasury has as yet only agreed to cover £2.5 million.

The SRC has a long way to go, and for the past three weeks has been seeing what can be trimmed, in a fashion which has put even more people on edge. Requests have gone out to directors of establishments to look at their capital expenditures with a view to making immediate savings. Major establishments will have to cut capital expenditure by a fifth in the next year and one round of new research grants, due in a month's time, is going to be deferred. The other research councils are also having to bear the load of CERN and ESA.

No one would pretend that all this unseemly grubbing around is going to bring Britain's science to a grinding halt within months. But it does seem to have generated a fair amount of unnecessary bad feeling amongst scientists. And it does pose problems about whether the whole structure of scientific research—big science and little science alike—should be so easily rocked by international exchange rates.

In the short term, it is incumbent on Mrs Williams, convinced European that she is, to go in fighting for automatic reimbursement of the increased European subscriptions. But in the longer run there is no doubt that some way must be found of giving the chairmen of research councils increased flexibility in their transactions. As it is they are committed every year down to their last penny and have no source of stand-by credit on which they could draw in situations such as these and which they could repay gradually if special pleading found no favour with the Treasury. □

● In *Nature* of September 30, page 359, the maximum permissible body burden for plutonium was given as 0.6 mg. This should, of course, have been 0.6 µg.

The Nobel prizes (1): Physics

Pioneers of the New Physics

Stuart Sharrock puts the achievement
of Richter and Ting into perspective

THE discovery of stable elementary
particles of high mass has opened up a
new era in experimental and theoretical
particle physics. For their pioneering
work in this discovery Professor Burton
Richter of the Stanford Linear
Accelerator Center (SLAC) and Profes-
sor Samuel Ting of the Massachusetts
Institute of Technology have been
jointly awarded the 1976 Nobel Prize
for physics.

The existence of a new particle was
announced simultaneously by the two
experimental teams in November 1974.
It was observed independently by
Ting's group working on the proton
accelerator at the Brookhaven National
Laboratory (BNL), and by Richter's
team at the electron–positron storage
ring facility at SLAC. The new
particle, called J at BNL and ψ at
SLAC, is a neutral meson of high
mass, 3.1 GeV, and with the amazingly
narrow width of 100 keV. It is difficult
to detect with proton machines, and its
discovery at BNL was a great feat of
experimentation. In contrast it is
copiously produced in electron–positron
annihilation, although its narrow width
makes it elusive. Within a few days
of the announcement of the J/ψ the
SLAC team had discovered a further
state, the ψ' at 3.7 GeV.

The existence of stable particles of
high mass was unexpected. The J/ψ
is a spin 1 particle and is more than
twenty times heavier than the lightest
such particle, the pion. The J/ψ should
rapidly decay into lighter known par-
ticles, yet its lifetime (related to its
width through the uncertainty prin-
ciple) is about 1,000 times longer than
expected. Furthermore the J/ψ does not
fit into the well-established three-quark
classification scheme. A new law of
nature is therefore necessary to explain
its decay properties and extremely long
lifetime.

Of the four known forces between
particles only the electromagnetic inter-
action is understood. Both Richter and
Ting have in the past contributed to
this understanding by demonstrating
the validity of the theory down to
distances of the order of 10^{-14} cm.
Attempts at a unified field theory for
the weak and electromagnetic inter-
actions based on the principle of 'gauge
invariance', in which an additional
symmetry is assumed to exist between
the two interactions, have had some
success in recent years, particularly
with the prediction and subsequent

observation of the neutral current weak
interactions. The theories require the
existence of a fourth quark, dis-
tinguished by the name 'charm'. This
new quantum number has successfully
explained the J/ψ particle as a bound
state of a charmed quark and a
charmed anti-quark.

If the charmed quark is heavy then
the long lifetime of the J/ψ is ex-
plained. The bound state of charm-
anticharm is similar to that of the e^+e^-
system in positronium, and is called
charmonium. A number of energy
levels are expected in this model and
many states have now been found fit-
ting well into the predicted level
scheme. The investigation of charmo-
nium promises to play a key role in
the understanding of strong inter-
actions by providing information on
the force acting between quarks.

Particles are classified as hadrons
if they have strong interactions and
leptons if they do not. The addition of
a charmed quark to the three conven-
tional quarks provides a set of building
blocks from which all known hadrons
can be constructed. It also provides a
symmetry between leptons and quarks,
and there are now four of each with
interactions that can be expressed in
terms of the same set of mathematical
rules. The detection of the new par-
ticles has given rise to a new branch of
hadron physics which could lead to a
better understanding of the symmetries
governing the structure of elementary
particles. Just as the discovery of
neutral currents may help to unify the
weak and electromagnetic interactions,
the concept of a charmed quark could
well indicate links between the weak
and strong interactions.

Professor Ting's experiment at BNL
was part of a systematic and precise
study of lepton pairs produced in
hadron–hadron collisions. This study
was started in 1972 and, to cover the
mass region 1–50 GeV, involved experi-
ments at BNL, the ISR in Geneva and
the electron synchrotron DESY in
Hamburg. One of the aims was to
search for long-lived particles decaying
into lepton pairs. These experiments
are very difficult as the rate of electro-
magnetic pair production in hadron–
hadron collisions is extremely low
compared to the rate of production of
hadron pairs. The experiment had to
reject hadron pairs by a factor of at
least 100 million. During the period
August–October 1974 the Ting group

Burton Richter

Samuel Ting

had seen a narrow peak in the effective
mass distribution of electron pairs, and
had performed many detailed checks to
prove that the peak was a real particle.
The announcement of the discovery
was withheld at first while the experi-
ment was modified to search for pos-
sible further new phenomena. The in-
evitable rumours began to circulate
and the group decided to publish at the
beginning of November. Around this
time the new particle was also seen at
SLAC and the results of the two
experiments were published simultan-
eously.

At SLAC, Richter's team had been
working with the immense magnetic
detector at the electron–positron stor-
age ring complex (SPEAR). They had
compiled a large volume of data on the
total cross-section for e^+e^- annihilation
as a function of energy. On checking
through these data an anomaly was
observed at one particular energy set-
ting: inconsistent values of the total
cross-section were obtained in measure-
ments taken at the same nominal
energy. The group decided this
anomaly had to be cleared up, and at
the beginning of November they
decided to investigate the energy
region in detail. Within hours of start-
ing an extremely narrow peak in the

cross-section had been seen. This effect was so strong and clear that the group immediately sat down and wrote the draft of their paper announcing the discovery of the new particle. The following day Richter and Ting met at a committee meeting and to their mutual astonishment started to tell each other about the interesting physics results they had found.

Burton Richter was born in 1931 in New York. He decided to become a scientist as a child in high school, a physicist as an undergraduate, and a particle physicist after obtaining his PhD from MIT in 1956. Richter, always fascinated by electrons, sought a job at SLAC where he performed experiments demonstrating the range of validity of quantum electrodynamics

down to very small distances. He has long been a proponent of the virtues of electron–electron scattering and was instrumental in the commissioning of the e^-e^- colliding beam machine at SLAC. During this time he designed an e^+e^- storage ring and worked for ten years to create the complex known as SPEAR. His team working on SPEAR began experiments in 1973 culminating in the discovery of the ψ family of particles and the subsequent measurement of their properties.

Samuel Ting was born in 1936 in Michigan but returned to China with his parents when he was two months old. At the age of 20 he returned to America with $100, very little knowledge of English and a determination to go to university. Within a short time

he had obtained a degree and a PhD from the University of Michigan and decided to become an experimental physicist. After a year at CERN he returned to teach at Columbia University in 1965. At Columbia he proposed an experiment at DESY to check a recent experimental result which appeared to show a violation of quantum electrodynamics. This was accepted, and in 1966 he began a long association with DESY in which he performed experiments extending the range of validity of quantum electrodynamics and determining the properties of vector mesons. His work over the last decade on lepton production culminated in the discovery of the J/ψ by its decay into electron pairs. □

The Nobel prizes (2): Chemistry

Boranes surgery

A **Special Correspondent** looks
at the work of William Lipscomb

IT is no surprise to chemists that Bill Lipscomb should have been awarded the Nobel Prize for his beautiful work on the boron hydrides. The layman may, however, be baffled, wondering what these oddly-named entities might be, and why they should have merited years of study by a leading crystallographer who is no less well known for his more recent work on the structure of enzymes.

The story goes back more than half a century, to the time when the great inorganic chemist Alfred Stock reported that the simplest compound of boron and hydrogen is not BH_3, as one would expect by analogy with BF_3, but its dimer B_2H_6. The trouble with B_2H_6, and with all the other boron hydrides which he described, was its "electron deficiency". If the molecule had an ethane-like geometry (as the evidence of electron diffraction was held to indicate) it would need 7 electron-pair bonds to hold its atoms together, but only 12 valency electrons were available for the purpose. This ominous crack in orthodox valency theory was papered over by Linus Pauling, who advanced special reasons why the B–H bond should be able to make do with less than 2 electrons, and proposed structures for the higher boron hydrides in line with this idea.

In 1943 Longuet-Higgins and Bell showed that the chemical and spectroscopic evidence on B_2H_6 harmonised less well with an ethane-like structure than with a bridge structure, in which

two of the H atoms are situated in the middle of the molecule. It appeared that B_2H_6 manages to hold together, not by starving all its bonds of electrons but by sharing two electron pairs between the two sides of the central bridge. The bridge structure was confirmed soon afterwards by W. C. Price, and in 1949 it was proposed that one should regard B_2H_6 as containing two "liaisons triatomiques", or "3-centre bonds", neatly describable in terms of localised molecular orbitals. But the higher boron hydrides remained a puzzle.

At this point the crystallographers grasped the initiative, and Kasper, Lucht and Harker led off with a definitive determination of the molecular structure of crystalline $B_{10}H_{14}$. Their results if anything deepened the mystery: the bonding in this molecule clearly involved principles more general than those which had been advanced for B_2H_6, and it became clear that a full understanding of these principles must await the collection of structural evidence on the other boron hydrides.

It was Lipscomb who took up this challenge. Undaunted by the pathological instability of most boron hydrides and the acute difficulty of "seeing" hydrogen atoms by X-ray diffraction, he and his colleagues completed a sequence of brilliant experimental studies of the hydrides described by Stock, and of others isolated in his own laboratory. A pattern soon

William Lipscomb

began to emerge which could be interpreted in terms of "3-centre bonds" of various novel kinds, together with "many-centre bonds", for whose possible existence there were solid grounds in molecular orbital theory. Simultaneous theoretical studies in England and by Lipscomb's group predicted the existence of a stable ion $B_{12}H_{12}{}^{--}$ with icosahedral symmetry. .

The preparation of the potassium salt, and the subsequent determination of its structure by Wunderlich and Lipscomb in 1960, set the final seal on the molecular orbital theory of the bonding in these remarkable substances. The fascinating work of Lipscomb and his colleagues is a story of the fruitful interplay between imaginative theory and skilled experiment, leading to a genuine "breakthrough" in our understanding of the forces which bind atoms together. No longer is there a dichotomy between the chemist's "covalent" bond and the "metallic" bond of the physicist: Lipscomb's work showed us that at least one element succeeds, in its hydrites, in forming both kinds of bond within the same molecular species. □

The Nobel prizes (3): Medicine

Antipodean . . .

Carleton Gajdusek's attainments
stretch beyond virology. **Cedric Mims**
outlines his work and assesses its significance.

C ARLETON Gajdusek came into medical virology in the early 1950s and he soon developed a particular interest in the haemorrhagic fevers of the USSR and the Far East. His work on virus infections took him to many other parts of the world including South America, the Middle East and a number of Pacific Islands. Not content merely to investigate the viruses infecting a given population he also made extensive studies of the behaviour, development and language of the people. In all but name he became an anthropologist. During the 1960s he could be counted on to talk with authority not only about the exotic diseases occurring in peoples in remote corners of the earth, but also about aspects of their physical and cultural anthropology. In 1959, at the National Institute of Health in Bethesda, Maryland, he became simultaneously director of a programme for studying child growth and development in primitive cultures and director of the laboratories for slow, latent and temperate virus infections.

It was in 1957, during an expedition to the highlands of New Guinea, that he encountered the disease kuru, and together with Vincent Zigas, the local physician who already knew the disease, he brought kuru to the attention of the rest of the world. Kuru was a chronic neurological condition, progressive and fatal, affecting especially the cerebellum and occurring in a sharply localised area inhabited by the Fore tribe. In these people it was responsible for more than half the total deaths after infancy, and in some villages at least half the women were affected.

Initial thoughts about the role of local genes, local plant or other toxins, were dramatically swept away after Dr Hadlow, of the Rocky Mountain Laboratory, Montana, had pointed out that the brains of patients with kuru bore a histological resemblance to the brains of animals suffering from scrapie. Scrapie, another chronic degenerative neurological condition affecting sheep, was known to be transmissible experimentally from animal to animal, but only after an unusually long incubation period of a year or two. Gajdusek therefore flew refrigerated brains from patients with kuru to the Washington laboratory where in August 1963 his close col-

laborator, Joe Gibbs, injected the brain suspensions into two chimpanzees. Twenty-one and thirty months later the injected animals were seen to sicken with a disease closely similar to kuru and their brains showed almost identical pathological changes.

The disease could be transmitted from chimpanzee to chimpanzee, in whom the average incubation period was 22 months, and in subsequent experiments various other primates were found to be susceptible. All the usual laboratory tests for microorganisms were negative, except for a number of chimpanzee viruses which were recovered and appeared to represent the normal viral flora of the brain. The transmissibility of the agent from animal to animal and its assay ($10^{7.5}$ infectious doses present per gram of brain) firmly established that kuru was caused by an infectious agent. But kuru, being closely similar in its properties to scrapie, was not a conventional virus and to this day its composition and mechanism of replication is unknown. For the first time therefore, a chronic neurological condition of man, in which there were no pathological or other signs of infection, had been shown to be due to a transmissible agent.

'It seems probable that we have still not gathered the full harvest of the discoveries made by Gadjusek'

Throughout this work and during his frequent visits to New Guinea, Gajdusek had both the opportunity and the energy to exercise his anthropological interests. As a result of this he wrote various monographs, produced films and contributed to paediatric journals. The marriage of virology and anthropology bore appropriate fruit when it became clear that the disease kuru in New Guinea was transmitted from person to person by cannibalism. The disappearance of cannibalism in this part of New Guinea has meant that kuru too has now almost disappeared.

Kuru is restricted to a small zone

in the highlands of New Guinea, where a total of some 2,000 cases has been reported since its discovery in 1957. But the discoveries about kuru have sparked off a fresh approach to other chronic degenerative neurological conditions of man which are of unknown aetiology, such as amyotrophic lateral sclerosis and parkinsonian dementia. So far one of these conditions, called Creutzfeld Jacob disease, has been shown in Gajdusek's laboratory to be transmissible to primates and caused by a replicating infectious agent similar to kuru. Creutzfeld Jacob disease is rare, but the infectious agent has now been recovered from the brains of patients in five continents of the world. It seems quite likely that the neurological disease is the tip of the iceberg and reflects widespread sub-clinical infection.

At about the same time as the discovery of kuru, certain conventional viruses had been shown by others to be responsible for obscure and chronic neurological conditions. Measles virus was recognised as a cause of the rare disease subacute sclerosing panencephalitis, and JC virus, a new human papovavirus, came to be associated with progressive multifocal leucoencephalopathy. Once again the infectious agent was seen to persist in the body and cause a disease after a period of several years.

It is against this background that Gajdusek's work on kuru must be assessed. It is now firmly established that chronic diseases of man, especially neurological diseases, can arise as a result of the slow but relentless multiplication of certain infectious agents, sometimes referred to as "slow viruses", which cause disease many years after initial infection. The possible incubation period with kuru, for instance, is between 10 and 15 years. Much of the recent excitement in multiple sclerosis research arises from observations based

Carleton Gajdusek

on this concept. It represents a great step forward in our understanding of disease and it seems probable that we have still not gathered the full harvest of the discoveries made by Gajdusek. These discoveries, moreover, may well prove to be of more than mere biological interest, because traditionally microbiology has moved from the recognition of an infectious aetiology to the development of a vaccine and thus to the prevention of the disease.

The agents responsible for kuru, Creutzfeld Jacob disease and scrapie pose major problems for the investigator. All assays and tests depend on the lengthy incubation period in experimental animals and it may take more than a year to learn the result of a single experiment. Other microbiologists would have shied away but Gajdusek's insight, patience and persistence have ensured that the infectious origin of kuru was discovered and the

disease given an important place in our understanding. He has been the pioneer and the central authority in this area of biomedical research. His encyclopaedic memory, his familiarity with strange places and primitive peoples have generated many legends. He is the sort of man one might have expected to win a Nobel prize and at the same time it is both a surprise and a pleasure to find such a gentle and likeable man behind the legends. □

... encounters

A **Special Correspondent** describes
the discovery made by Baruch Blumberg,
'one of the success stories of the century'

SURPRISE, suspense and success are the three indispensable ingredients cin a good detective story, and none of them is absent from the events which led up to the discovery and evaluation of Australia antigen, now usually known as hepatitis B antigen, or HB antigen.

The story starts nearly 20 years ago, when Dr Baruch Blumberg, one of the recipients of the 1976 Nobel Prize for Medicine, was at work on proteins in the blood. It is well known that humans show individual variation in the antigens, or blood group substances, on their red blood cells. Blumberg set out to detect and compare comparable substances in the plasma as opposed to those on the cells. Using as antisera specimens from patients who had received many transfusions, a reacting antigen was found in an Australian aborigine.

The surprises began when this was found to be widely, if unevenly, distributed in human blood throughout the world. The United States, where Dr Blumberg was working, was in fact relatively free from it, although a survey of patients revealed that cases of leukaemia, Down's syndrome (mongolism) and hepatitis figured among those who were positive. The question began to be asked: was this in fact not a simple protein antigen but a human leukaemia virus?

The possibility of being able to detect and study such a virus opened wide horizons—but it was now that the suspense began, because it quickly became unlikely that it was, in fact, a leukaemia virus. A common factor in leukaemia and Down's syndrome is altered cellular immunity; in addition, patients with leukaemia have many transfusions with a concomitant oppor-

tunity for acquiring hepatitis, and children with Down's syndrome (at least, those of them in institutions) are notoriously prone to hepatitis.

It was about this time that hepatitis B infection, acquired from transfused blood, broke upon the haemodialysis trade. The transmission of hepatitis B virus, undetectable by any means except the production of the disease in a patient, had always been a problem in blood transfusion and the use of blood products, but the disasters which overtook staff and patients in dialysis units had made it doubly urgent to find a means of recognition, a serological marker for the presence of the agent.

It is now well known how, during the late 1960s, it became clear that Australia antigen was such a marker—but was it, in fact, the hepatitis B virus? When examined by negative staining it did, so to speak, come to life, and at least some of the particles could be seen to be acceptable as virus. By now it has been possible to purify these particles, and to determine that they have double-stranded DNA and a DNA polymerase, and in fact to do many things with them except grow them in a laboratory system.

Many medical microbiologists must have hoped that this would give a clue to growing hepatitis B virus in a cell culture system, but so far they have been disappointed. However, this has not prevented the use of HB antigen to study the natural history of hepatitis B virus, to detect carriers, and, to detect antibodies in human serum, and this is where the third ingredient, a successful conclusion, comes in.

As a corollary, it is possible to prevent many cases of hepatitis B (post-transfusion and otherwise) and hence to make haemodialysis, blood transfusion,

Baruch Blumberg

and the use of blood products infinitely safer. Unfortunately, there is still no conventional small laboratory animal, but the chimpanzee serves as a possible, if expensive, source and host. Even a vaccine has been prepared from HB-antigen-positive human blood, and even though its worth has yet to be assessed, there is little doubt about the effectiveness of the passive immunity conferred by the injection of concentrated, purified, specific anti-HB immunoglobulin.

Although the final evaluation of HB antigen involved, in the end, many laboratories, the work all springs from the researches of one principal and primary investigator. Even if, by hindsight, the connection between the molecular genetics of polymorphism and the prevention of hepatitis B in haemodialysis units all over the world seems to involve a series of knight's moves in chess, the consequences of Blumberg's discovery, judged as a piece of preventive medicine, must rank as one of the success stories of the century. Perhaps in this respect it is as near as any previous award to Alfred Nobel's desire to reward those whose discoveries have been of benefit to mankind. □

Now New York steps in

The New York State Attorney General's office is considering whether or not to impose controls on recombinant DNA experiments performed in facilities in the State. Colin Norman reports from New York

A windowless hearing room on the 47th floor of New York City's World Trade Center last week provided a somewhat unlikely venue for another round of public hearings on the potential risks and benefits associated with recent advances in genetic engineering. The hearing marks a potentially important development in the long process of regulating recombinant DNA research in the United States, for it is the first time that the matter has been taken up by a state government.

Called by the New York State Attorney General, Louis J. Lefkowitz, the hearing was designed to elicit the views of various concerned individuals —mostly scientists who have been in the thick of the debate over the potential hazards of recombinant DNA research for the past two years—on whether New York should impose its own controls on the research and, if so, just how strict those controls should be.

The crux of the discussion was concerned with two closely related issues. First, whether the guidelines, published by the National Institutes of Health last June to control recombinant DNA experiments which NIH supports, are strict enough to guard against possible public health hazards. And second, whether or not the guidelines should be enforced in New York State by law.

The outcome of the discussions in New York will be closely watched elsewhere in the United States, for other states are likely to take up the matter at some stage. Already, the City Council of Cambridge, Massachusetts, has imposed a temporary moratorium on a few types of recombinant DNA experiments at Harvard and Massachusetts Institute of Technology, while a special review board, consisting of city residents, is looking into the risks and benefits associated with the research. (The Cambridge moratorium was due to expire on October 7, but it has quietly been extended until January 7 to allow the board more time to make its recommendations.) And a committee of the City Council in San Diego has also

been taking evidence from both supporters and opponents of the research. New York is the first state to consider adopting legal controls, however.

According to one official in the Attorney General's office, a major factor which led New York to look into the matter is the fact that the NIH guidelines apply only to NIH grantees, and there are no legal sanctions which can be applied if the guidelines are violated. Research supported by industry, and even by other government agencies, is not formally covered. Thus a prime consideration in New York is whether or not the controls should be backed by the force of law and made to apply to all facilities, including industrial laboratories.

Last week's hearings consisted of a parade of about a score of witnesses who urged the State government to undertake a variety of measures ranging from doing nothing to banning virtually all recombinant DNA experiments in the state. The hearings were mostly low-key, punctuated by some fairly gentle questioning by officials in the Attorney General's office. They contrasted rather starkly with the highly charged, emotional nature of the Cambridge City Council's hearings earlier this year.

Leading the call for a moratorium on the research were George Wald of Harvard, Leibe Cavalieri of the Sloan-Kettering Institute. Erwin Chargaff of Columbia University and Jonathan King of Massachusetts Institute of Technology. Cavalieri kicked off with a critique of the guidelines and the way in which they had been developed, arguing that there had been only "token" input from the public, and suggesting that "it is up to society to see that the research effort is directed toward solution of problems, not the creation of new problems". Cavalieri was the only speaker to dwell at length on what he called "the hazards of success", suggesting that so far the discussion of recombinant DNA research has failed to consider the implications of the successful manipulation of the genetics of, for example, crop plants. The result, he said, "would be a change in the intricate balance of our environment, which has taken millions of years to establish".

Wald was the most outspoken critic of the research and the guidelines, however. He suggested that the recent outbreak of Legionnaires

Disease in Philadelphia has provided "a beautiful model" of what a public health problem with recombinant DNA research would be like— "unidentifiable and impossible to trace to its source". Wald argued that recombinant DNA research should not be conducted in universities, but it should be isolated in one or two national safety facilities.

As for proponents of the research, the most outspoken was James Watson, Director of the Cold Spring Harbor Laboratory. Watson said that he once told former Presidential candidate Sargent Shriver that the furore over recombinant DNA research is "the most overblown thing to enter the American scene since (President Kennedy) created the fallout shelter". Watson argued he could see no danger in transplanting genes from higher organisms into viruses or bacteria, and he said that he could see little hazard even in transferring genes from viruses known to cause cancer in animals into bacteria. The impact of infection by such organisms would be "negligible" compared with the impact of infection by the viruses themselves, he said, and suggested that "the marginal danger of this thing is a joke compared with other dangers". He suggested that the Attorney General's office could better spend its time looking into the potential health hazards of hair dyes and flame retardants in children's clothing. "What started as an attempt by the scientific community to be responsible takes on the aspects increasingly of a black comedy", Watson stated.

Faced with such widely conflicting viewpoints, the New York Attorney General's office is going to have a tough time deciding whether or not to impose additional controls on researchers in New York. Its decision will, however, be carefully watched in other state capitals. □

● **Eleanor Lawrence adds from London:** The US National Institutes of Health guidelines and the report of the British Williams Committee have now been followed by recommendations for the conduct of *in vitro* recombinant DNA research by two European organisations, the European Science Foundation (ESF) and the European Molecular Biology Organisation (EMBO).

The ESF makes a number of suggestions. It says national registries of recombinant DNA research, analogous to the Genetic Manipulation Advisory Group proposed for Britain, should be set up in each country. It also suggests that all industrial, university and government laboratories be legally obliged to declare relevant aspects of

their work to such a registry. The supervisory and monitoring machinery would remain a national responsibility. There is a further recommendation that the code of practice drawn up by the Williams committee be followed rather than the NIH code since Williams specifies a greater degree of physical containment for most experiments, and is more flexible.

The EMBO guidelines also support the idea of national advisory groups, which would specify the conditions of containment to be used. EMBO says that either the Williams or the NIH codes of practice may be followed, but not a mixture of both. The different degrees of emphasis on biological and physical containment in the two codes, it contends, could lead to confusion.

Both organisations emphasise the need for regulations to operate at the same level in different countries and

therefore suggest that the national advisory committees should meet regularly. In addition they both propose that suitable experiments should be undertaken to assess the conjectural risks associated with recombinant DNAs and to pave the way towards eventual adjustments to the guidelines.

In Britain, however, many biologists are unhappy about the draft Health and Safety Commission (HSC) regulations now being circulated for comment. The regulations require among other things compulsory notification to the HSC of any experiment intended "to alter or likely to alter the constitution of any microorganism". Although the HSC has asked which experiments should be excluded, this type of blanket regulation, and the fear that proposals submitted to the HSC would be delayed interminably by bureaucratic machinery, have stimulated a group of

scientists to ask government that the HSC regulations should be withdrawn completely. They wish to see the voluntary procedures recommended by Williams followed instead.

There is strong feeling that in this instance the HSC is not the most appropriate body to deal with the complex scientific issues involved and that public interest would be better served if they were considered by an expert committee such as the Genetic Advisory Manipulation Group for example. The Health and Safety Commission would still have powers under the present Health and Safety at Work Act to inspect laboratories to see whether the containment procedures recommended by Williams were being implemented, and could probably be brought in by GMAG to prevent experiments considered unsafe being carried out. □

COMECON

● Scientific cooperation between Comecon countries is still "incommensurable" with the requirements and potentialities of the member countries, according to the Polish journal *Sprawy Miedzynarodowe*. The relatively slow development of joint research and development projects is to a certain degree attributed to "subjective factors" such as lack of personal initiative, but also to the "excessive autarchic tendencies" of individual bodies which often necessitate intervention by the "central authorities" (that is, Party and Government) of the member countries.

A number of remedies are proposed to help the integration of research which is a fundamental tenet of Comecon planning. Most of these relate to improvements in organisation and management, the integration of science and industry and so on. One, however, suggests a basic lack in the whole policy of cooperation. This involves the question of funding. At present, it appears, the necessary funds are obtained often only after protracted negotiations between the countries concerned. To speed up the process, it is suggested that a special fund should be created within the International Bank (in Moscow) for the integration of research and development in the Comecon countries, with, in due course, the possible establishment of a Comecon bank especially for the financing of research.

● Since the discovery of lasers, laser research has been seen as a modern and progressive subject, eminently suitable for Socialist countries. The Second International Conference on

Luminescence at Szeged in Hungary last month showed that interest continues unabated. In particular, a Hungarian team described the use of lasers to determine "hitherto-undiscovered properties of substances", while Soviet, Polish and East German specialists reported considerable success in the use of lasers to investigate photosynthesis.

● Cosmonauts from all the Comecon countries—including Cuba and Mongolia—are to take part in manned space-flights between 1978 and 1983 as part of the Interkosmos programme. Until now, the programme has consisted solely of the launching of unmanned satellites, put into orbit by Soviet rockets and carrying experiments and equipment provided mainly by the Soviet Union with some participation from the European members of Comecon, notably East Germany and Czechoslovakia. Not surprisingly, Soviet participation in the new programme will also be considerable: as well as providing the spacecraft, orbiting stations, and launch facilities,

the Soviet Union will train the cosmonauts—training will take place at the Soviet "Yuri Gagarin" training centre. And the wording of a recent TASS communique suggests that at least one crew member in each flight will be from the Soviet Union.

●*Chemicisation*, the massive, widespread, undesirable use of artificial fertilisers, has long been a cornerstone of Soviet agricultural policy, and, not surprisingly, has become standard practice in other Comecon countries. In certain cases, however, the results have not been entirely successful. In Hungary, for example, excessive use of fertilisers has considerably reduced the sugar content of sugar beet in recent years, so that while the yield of beet itself has increased steadily, the overall sugar yield has remained constant. To encourage the production of maximum-yield crops rather than overall gross bulk, the Director-General of the Hungarian Sugar industry recently announced a new scheme of payment for sugar-beet crops under which farmers will receive a premium based on sugar content. The Hungarian experience is, apparently, not unique: recent Soviet resolutions on the further improvement of agriculture stress the need "to increase research on the study of the effect of mineral fertilisers and other chemical products on the quality and biological total value of agricultural production". This marks a radical change from traditional attitudes: until now the largest crops rather than the best have always won official acclaim.

Vera Rich

Same song, different meaning

Chris Sherwell reports on the latest developments on the European Community's research front

THE EEC's Council of Research Ministers did their turn again at the end of last week when they met in Luxembourg. In most respects it was a repeat performance of a long-running drama, but it did have the added merit of a new act to take it further towards the worthy ending it has so far lacked.

The Ministers were unable to make the crucial and much-delayed decision on the siting of the Joint European Torus (JET), the Community's sorely afflicted fusion project. It was that which made the meeting a repeat performance. But earlier expectations of progress on this particular matter had begun to die anyway, principally because the whole issue has in recent months become inextricably intertwined with the Community's four-year joint research programme, now under discussion.

When the Ministers left Luxembourg they had at least resolved a good deal concerning JET apart from the site. They were able to agree certain details regarding financial arrangements (though not others), as well as the project's legal status. It was these accords which helped mainly to redeem the show, but until the site is decided they can have little real practical meaning.

The meeting agreed that the European Commission will put up 80% of the costs, with the host country bearing 10% and the other participating countries the remaining 10%. The host country will also meet the costs associated with supplying the site with basics and with dismantling the project on completion in several years' time. At that point the equipment will be the host country's property. The Ministers further agreed that the project was a "joint enterprise", that is, independent of the host country and a Community project to be identified as such.

They were, however, unable to decide on staff numbers, although there was agreement that the host country would second people to the project and that others would be engaged as temporary agents of the Commission. Outstanding matters will be discussed by the member states' permanent representatives in Brussels, as will another issue, the precise costs involved. Since the Commission contribution of 108 million units of account (at March 1975 prices) was agreed, costs have risen and the project is being talked of in terms

of a four-year construction effort rather than five years. Some modification may therefore be necessary.

Much of this was largely in line with Commission proposals made a couple of weeks earlier. Regarding the site, however, the Commission was in favour of the Joint Research Centre (JRC) at Ispra in Italy—itself nothing new—but had made a new proposal should no Council decision be reached before the end of the year. In that event, it suggested, the matter should be passed over to the Commission itself, which would then decide on the basis of an opinion from its own JET committee— which decided matters by a two-thirds majority.

This, a precedent which did not look like attracting many Ministers, was not discussed in Luxembourg. Another complication had already crept in some weeks earlier, when it was agreed to discuss JET in the context of the four-year joint research programme. Although designed to hasten agreement on the JET site, it introduced a bargaining element well suited to Community *realpolitik*. With details of the research programme themselves matter of dispute, movement on JET became contingent upon movement on the research programme.

At last week's meeting demands concerning the programme came, perhaps ironically, from three of the four main contenders for JET— Britain, France and Germany. Britain's starting position, on which it is flexible, was to seek reductions in staff of about 20% and cuts in certain expenditures amounting to £15 million at the Community's four research centres. Italy argued that this would threaten the centres, especially Ispra, the fourth JET contender. No one is saying that Italy and the Commission are now preparing for the worst, but there is speculation that if JET isn't to go to Ispra only a large alternative research programme—perhaps the one proposed, intact—would provide sufficient compensation.

Certainly the British were keen to emphasise that the site decision ought to be taken on scientific and technical grounds only. That would help the Culham site's case, but there would still be the matter of the French and German sites of Cadarache and Garching. Of the two, Germany with its stronger economy is the stronger candidate, yet the opportunity to strengthen its case still furhter through a higher host country contribution was apparently allowed to pass. French

research faces similar problems to Britain's, but this may say more about the relative weakness of its candidacy than Britain's ability to take on the project. The latter's enthusiasm for the project going to Culham remains undiluted by the current crisis facing its research community and economy generally.

As the Luxembourg meeting broke up the recipe was there for a good deal of haggling between the parties between now and the next Research Ministers' meeting on November 18. The understandings reached on JET and the joint research programme are provisional until a decision on the JET site is reached. If that decision, already delayed a year, is not made then, it could pass directly to the Council of Prime Ministers. If they didn't agree and refused to leave the matter to others, the Commission could go on funding the project in the way it has done hitherto. The fate of the project, however, will be more uncertain than ever. □

● The European Science Foundation (ESF), members of which are national research councils and academies in Europe, is due to hold its general assembly in Strasbourg this week. The Foundation involves a total of 43 member organisations from sixteen European countries. The President of the Executive Council, which will present its second report to the assembly, is Sir Brian Flowers, the UK physicist.

The Foundation aims to encourage the emergence of a close-knit community of science and research in Europe; this is also a goal of the EEC proper, with which the ESF has links through the Director General for Research, Science and Education at the European Commission. The ESF endeavours to do this by providing a forum for the exchange of information and experience, particularly regarding the way the member organisations actually operate. It also seeks by this method to encourage activity aimed at producing joint European action in certain fields.

As an international non-governmental non-profit organisation it lacks the financial means to give effect to its advice, but it can and does examine specific areas and maintains contacts with other concerned organisations, like EMBO and ESA. The Executive Council reports further progress in such areas as astronomy, space science, genetic manipulation and mathematics, in all of which it had sought to take initiatives during its first year of existence.

IN BRIEF

SERI delay

Another three-month delay has been announced by the Energy Research and Development Administration (ERDA) in the timetable for establishing the Solar Energy Research Institute (SERI), the facility which may eventually become the focal point for solar energy research in the United States. ERDA was hoping to choose a contractor to operate SERI and a site for the facility in December, but it now says that a decision cannot be made before March.

The delay stems chiefly from the fact that there is fierce competition among states for the contract to operate SERI, which will bring considerable amounts of federal money, jobs and prestige. ERDA has received 20 proposals for the contract and so far has eliminated only one of them.

The three-month delay is only the latest in a series of setbacks to the plans for SERI, first approved by Congress in November 1974.

SGHWR report shortly

Last week the final witnesses gave evidence on Britain's choice of nuclear reactor to the House of Commons Select Committee on Science and Technology. The committee says it hopes to publish its report in about a month, having questioned the Energy Minister Mr Benn, the Central Electricity Generating Board and the South of Scotland Electricity Board, the UK Atomic Energy Authority and the National Nuclear Corporation, and the Health and Safety Executive and Electrical Power Engineers Association.

Since the choice of the British Steam-Generating Heavy Water Reactor two years ago it has undergone substantial modification (a reference design was only submitted in June of this year) and expenditure on it was deferred for a year in this summer's government spending cuts. Continuing doubts about whether it was the best choice made the committee decide to look at the whole question afresh.

Protest over Nacekuke closure

Civil service unions are protesting to the Secretary of State for Defence about the plans to close the Chemical Defence Processing Establishment at Nancekuke, Cornwall—due, according to the unions, in March 1978. Some of the work of the station will be moved to Porton Down in Wiltshire, another defence establishment.

The union objects on the grounds that Nancekuke's function is the production of defence chemicals, a job that Porton Down, a primarily research establishment, is not equipped to do. According to the 1973 Nugent report, closing the station would save around £80,000 a year, mainly in administrative costs, but the unions claim that money would not be saved.

LEGISLATION to control toxic chemicals was as inescapable as are the facts that all substances are chemicals, all chemicals (depending on the level) are toxic, and all human activities are the target of regulation. Passage of the new law was accelerated by some notable gaffes by industry. Some of these mistakes were inadvertent because they preceded recent advances in toxicological testing. In other cases there was a reluctance to adopt known procedures in testing industrial compounds; some of these had long been obligatory for proposed new agricultural chemicals and food additives.

Insecticides bore the brunt of the earlier demands for increased regulation, perhaps because the methods used in applying insecticides were often easily visible, as when residential areas were sprayed to control mosquitoes. Conversely, food additives got into trouble with the public partly by being honest. Their names are required on labels of foods in which they are put. Few scientists, and certainly no non-scientists, can derive a sense of epicurean satisfaction from contemplating the words "butylated hydroxytoluene" on the list of goodies in a prepared snack. Such nomenclature heightens the conviction that we are surrounded by a sea of new synthetic chemicals even though our species has been sniffing and enjoying the odours of allyl isothiocyanate and diacetyl for many generations.

The new law should bring protection to workers in chemical factories. Vinyl chloride, kepone and dichloromethyl ether furnished three examples of inexcusable incidents in which such workers were exposed

Curbing chemicals

THOMAS H. JUKES

to toxic amounts of injurious chemicals. There is uncertainty and disagreement as to where the levels of such substances become tolerable, especially in the case of known or suspected carcinogens. However, it is known that the time of onset of cancer is lengthened by decreasing the dosage level, and lowering the dosage also decreases the probability that cancer will actually occur. These relationships apply regardless of whether or not an actual threshold can be established.

When the pressures for bans are examined and compared there are some interesting paradoxes. Polychlorinated biphenyls (PCBs) escaped scot-free for about eight years after it had been shown that they were environmental contaminants on a global scale, toxic at low levels to many vertebrates, and distributed without control. Their American manufacturer made small gestures such as taking PCBs out of "carbonless carbon paper". Their use continued in transformers and capacitors, which furnish, among other things, a comfortable living for the middle classes. But the fury of middle-class environmentalists descended on DDT, the use of which was primarily needed to save lives and prevent disease in disadvantaged countries. PCBs are far more stable than DDT and are known to be injurious to human beings as contaminants of cooking oil. Fortunately PCBs are on their way out, after being used as examples for the need of the new bill.

Let us hope that the aroused consciousness of possible injury from "new chemicals" will eventually result in a balanced study of hazards from both synthetic and "natural" compounds. In the meantime I am trying to avoid a feeling that I am endangering myself when I sniff the smoke of a camp fire.

correspondence

Exotic viruses

SIR,—In last week's article on exotic viruses (page 625), an error was inadvertently introduced during the editorial work.

The identification of a Marburg-like agent announced two weeks ago was from material obtained from patients in a recent outbreak of illness in Zaire and the south of Sudan only.

ARIE ZUCKERMAN
*WHO Collaborating Centre for
  Reference Research
  on Viral Hepatitis,
London School of Hygiene
  and Tropical Medicine,
London, UK*

Engineers' action

SIR,—In your Editorial of October 14 on the Institution of Chemical Engineers' report on Materials and Energy Resources you take that body to task for attempting "both to stimulate reasoned discussion on crucial resource problems" and trying "to get something done about them." But the professional engineer's job is precisely "to get things done", not simply to establish the truth; and planning ahead is part of his work. So, when a body of engineers, having closely surveyed a field in which they have specialised knowledge and competence, foresee a dangerous situation developing, it is their duty not merely to point this out but to make recommendations for action—in this case by the government. And if not they—who else?

You suggest, too, that in such long-term matters "the Working Party might have done better to look more closely for international solutions". But practically-minded engineers are more humbly concerned with what can be done here and now, even when

their sights are on the future; and their chances of influencing government thinking, upon which you set such a low value, are at least a great deal higher than the likelihood of setting the world to rights.

P. M. C. LACEY
*University of Exeter,
Exeter, UK*

Whale resources

SIR,—Dr Robert May (Sept. 9, page 91) gave an interesting commentary on the problems of managing whale resources vis-à-vis other items in the the relative values of the discount rate and the potential rate of increase of the resource is stressed. In this comparison the discount rate can be defined as δ where the present value of a whale caught in x years time is $e^{-\delta x}$ times the value of a whale caught today. May implies that δ is approximately equal to 'a', the interest to be expected from a deposit in a bank (10–20%).

In an age of inflation, and changing values of products from limited natural resources vis-à-vis other items in the economy, this may be misleading. The value in pounds of a whale at the time it is caught in x years time can be expressed as $e^{(b+c)x}$ times the present value, where b is the general inflation rate and c the rate at which the net value of products from a whale has changed differentially relative to average values of all commodities. b is positive and c may also be if prices for meat and other products from whales increase faster, and/or the price of fuel and other resources used in catching whales increase slower than prices in general.

The value of the discount rate to be used in determining the economic attractiveness (net present value) of a

whale caught later rather than now is given by $\delta = a - b - c$, which may not greatly exceed and may even be less than the rate of increase of whale stocks of around 5% annually. That is, to the extent that δ is less than a, the economic self-interests of a well-informed sole owner of a whale resource will conflict less seriously with conservation policies than implied by May.

The more important reasons for poor whale management have been the lack of an owner and poor information. While whales are common property, the individual whaling enterprise (company or even country) can ascribe little value to a whale in the future since there is no guarantee that it will be able to harvest that whale or its offspring. Also in the absence of well-funded research, whaling interests have been poorly informed of the longer term implications of management proposals. In the resulting fog of uncertainty about the future, their interests have naturally focused on the present and more clearly visible events. Both these effects have resulted in undue emphasis being given to maintaining current catches at a high level.

On another point, it seems unlikely, given the low (less than 10% per annum) rates of mortality and reproduction of whales that the unexpected changes in catch rates of fin whales mentioned by May is really indicative of population instability. It seems much more likely that with the small number of whaling expeditions now operating, the catch per unit effort is a poor and highly variable estimator of the abundance of whales in any particular area in a single season.

J. A. GULLAND
*Department of Fisheries,
FAO, Rome*

Competition 10
Very Many Hairy Little Pigs Live In The Torrid Argentine = valine, methionine, histidine, leucine, proline, lysine, isoleucine, threonine, tryptophan, arginine (the amino acids essential to the rat). £10 for the best mnemonic in any field; entries by December 1 please.

Competition 9 asked for new words to describe some existing or future scientific concept or phenomenon. Entries galore: the winner, D. R. Reed of The Hague, The Netherlands, submitted six definitions, four of which we thought particularly successful: *cannobelism*— extreme form of competitive behaviour for scientific awards; *fornacation*—implantation of genetic (RNA based) information in human ova; *multiplication*—means of obtaining the product of two numbers without the use of remote microprocessors; and *plutophagy*—simultaneous solution of world population and radioactive pollution problems.

Honourable mention for R. A. Davis of Epsom, Surrey, with "*qualms*—any new elementary particles, in search of a name"; and Allan N. Zacher III of the University of Missouri, with "*vitrogen*—an adjective describing a substance produced in a cell-free preparation".

news and views

Destruction of stratospheric ozone?

from Michael A. A. Clyne

STRATOSPHERIC ozone will eventually be depleted significantly if releases of chlorofluoromethanes (CFM) continue at the present rates. Legislation is not necessary now, but if after no more than two years, existing evidence still supports significant ozone depletion, action to restrict manufacture should be taken by governments. These are among the main recommendations of the Committee on Impacts of Stratospheric Change, (*Halocarbons: Environmental effects of chlorofluoromethane release*, National Academy of Sciences, Washington DC, 1976) set up by the US National Academy of Sciences.

In 1974 I reviewed Molina and Rowland's article (*Nature*, **249**, 810) in the same issue of Nature, on the effect of chlorofluoromethanes on stratospheric ozone. Then, neither I nor most other scientists, believed that the hypothesis presented by Molina and Rowland, and independently by Stolarski and Cicerone (*Canad. J. Chem.*, **52**, 1610; 1974) would in fact develop into one of the most important recent issues in environmental pollution. Briefly, this theory, which has now largely been supported by an authoritative report of the US National Academy of Sciences (*Halocarbons: Effects on stratospheric ozone, Panel on Atmospheric Chemistry*, National Academy of Sciences, Washington DC, 1976), is as follows. CFM 11 (CFCl₃) and 12 (CF₂Cl₂), which contain chlorine, fluorine and carbon only, are produced in large amounts in all developed countries of the world. Approximately half of the total output is used as propellants in aerosol cans for

personal consumer uses (such as hairsprays and deodorants). Because of the stability that makes these compounds suitable for aerosol cans and for other uses, CFM 11 and 12 are very inert and no known major sinks for these compounds have yet been identified in the lower atmosphere (troposphere). Hence a large fraction of all past production of CFM 11 and 12, whether used for aerosols, car air conditioners or refrigerators is still present in the troposphere and indeed, these compounds will continue to accumulate there. Diffusion to higher altitudes occurs (Fig. 1), up to the stratosphere where most of the Earth's ozone (O₃) shield is located. The CFM can be photochemically dissociated there by far ultraviolet light at wavelengths around 190 nm. This light is not completely removed at stratospheric altitudes by the ozone and oxygen that normally filters out ultraviolet light, at wavelengths less than 300 nm, from reaching the Earth's surface. Photochemical action releases chlorine atoms from the CFM: and chlorine atoms react rapidly with ozone, giving chlorine monoxide (ClO) free radicals: $Cl + O_3 \rightarrow ClO + O_2$ .. (1). The reaction (1) is then followed at once by the rapid reaction (2) of ClO with the large quantities of the free oxygen atoms present at stratospheric altitudes between 25 and 50 km: $ClO + O \rightarrow Cl + O_2$ (2). Chlorine atoms are thus recycled, and are able to consume more ozone: this sequence is termed a chain reaction. Finally, the cycle is terminated by removal of chlorine atoms in their

reaction with methane

$$Cl + CH_4 \rightarrow HCl + CH_3 \qquad (3)$$

Most of the HCl formed in reaction (3) is not recycled back to active chlorine. Although many other reactions are involved, some of which are mentioned below, the overall effect of CFM is thus clearly to deplete stratospheric ozone. The NAS reports conclude that this description of the problem, essentially as advanced by Molina and Rowland, cannot be seriously questioned, although there are new factors that probably somewhat diminish the effect on ozone depletion that they calculated.

There is no justification for the total rejection of an effect of CFM on ozone depletion, which had been maintained earlier by a few scientists. In their report, the NAS Committee state that "Selective regulation of CFM uses and releases is almost certain to be necessary at some time and to some degree of completeness". In reaching their conclusions, the NAS were guided by a panel of distinguished experts including one Briton, Brian Thrush, and one German, Dieter Ehhalt. The report is an impressive document, containing a large amount of carefully sifted information that will be invaluable to future workers in atmospheric sciences. The approach is low-key, thoughtful and critical. The evidence accumulated by the panel is a formidable case in favour of an effect of halocarbons on stratospheric ozone. However, it is considered that the risks of a delay of not more than two years in restricting CFM releases are low. Although most scientists will probably agree with these views, some will doubtless point to the admitted wide margin of error in the estimate of a 7% eventual ozone depletion, with CFM 11 and 12 releases at the 1973 level. (According to the Panel, the limits of 2 to 20% ozone depletion have a 95% confidence level).

Some will say that the risk is slight, and that in any case industry is already reducing production of CFM 11 and 12, and is actively seeking substitutes that will have a shorter tropospheric life, and hence be less harmful. It re-
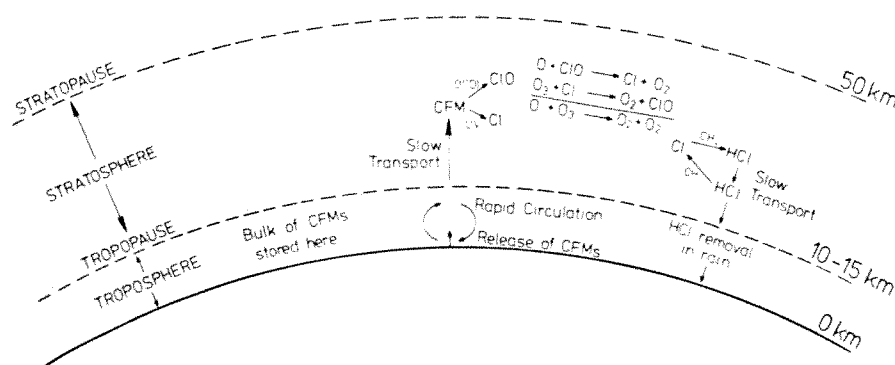
Fig. 1 The atmospheric ClOx cycle.

mains to be seen whether the search for acceptable substitutes for CFM 11 and 12 is sucessful; this is not likely to be an easy problem. CFM 22 (CHF₂Cl), among others, has promise as a substitute for CFM 11 and 12 in spray cans.

Other people will maintain that the risk is unacceptable, and that legislation is needed now rather than (possibly) in two years time. They will point to the possibility that delay might lead to politicians and governments losing sight of a problem which is effectively postponed for two years, a time scale comparable to intervals between elections. However, it should be noted that as well as studies by US government authorities, both the EEC and the West German government are initiating substantial investigations into the CFM-ozone problem.

The British government published their report on chlorofluoromethanes and their effect on stratospheric ozone earlier in the summer. This report was prepared by the Department of the Environment (DOE) with considerable assistance and guidance from AERE, Harwell. Their conclusions are broadly in line with those of the NAS report, although they do not recommend any specific regulatory action by governments. The DOE report is a useful summary of results; although clearly the resources that are evident in the NAS report were not available to the DOE compilers. The referencing of the DOE report is rather random and non-specific, and some literature citations are an encouragement to whimsical day-dreaming! For instance: ICI (1976), Imperial Chemical Industries, private communication. Dubreuilh W. (1896), des hypér-keratoses circonscrites; *Ann. Dermatol. Syphilig.* 3rd Series, 7, 1158-1204; quoted by Blum (1956).

In order to predict the effects of CFM on stratospheric ozone, a wide variety of data are required. The scope and uncertainties of these data varies enormously. Some of the basic types of information needed are (1) direct measurements of Cl atom, ClO radical, and other concentrations in the stratosphere; (2) the reaction rates of key elementary chemical reactions, as determined in the laboratory, and applicable to stratospheric temperature ($\simeq$ 220 K); (3) knowledge of the transport of air masses from the troposphere to the stratosphere. The first type of data (1), is perhaps the most important, because direct measurement of concentrations is the final check that the correct information under categories (2) and (3) is being used. Unfortunately, data of type (1) are by far the most difficult to obtain and are extremely sparse. The only results (so far) are *in situ* measurements of Cl and ClO

How rigid the nucleotide?

from a Correspondent

CRYSTALLOGRAPHERS have become used to seeing every structural idiosyncrasy they unearth in the crystal structure analysis of molecules of biological interest become a peg upon which the edifice of some molecular biological theory is hung. This has been particularly true for the study of the nucleotide and amino acid monomers from which nucleic acids and proteins are constructed. Beginning with the use of the covalent bond lengths and angles observed in monomers in the design of molecular models of these biological polymers, the results of crystal structure determinations have been extended to provide constraints which might properly be assumed for hydrogen bond geometry, for the conformations which can be expected to be preferred about various single bonds and, for the patterns of arrangement of water molecules and ions around biological polymers. This fertile field has been exploited with ever increasing confidence and perhaps nowhere more than in the analysis of the relationship between the properties and structure of nucleic acids. An important development in this work has been the rigid nucleotide concept

which suggests that mononucleotides have a rather rigid conformation which is conserved when they are joined together to form a polymer. X-ray diffraction studies of nucleotides, nucleosides and crystalline polynucleotides have provided a great deal of evidence for this although a number of recent X-ray single crystal studies of nucleotides have suggested that the situation may well be rather more complicated than was at first thought. The need for caution in applying the rigid nucleotide concept too insensitively is further emphasised by Evans and Sarma in a recent issue of *Nature* (**263**, 567; 1976). From NMR studies they suggest that adenine-containing nucleotides and the oligomers derived from them are no more rigid than the nucleoside adenosine itself. These studies are of course of particular interest because they relate to the situation in solution. An important application of the results of this work which is considered by the authors lies in the analysis of the structure of yeast phenyl alanine tRNA determined by X-ray single crystal studies and in particular the correlation of nucleotide conformation and base-stacking.

in the stratosphere from two balloon flights carried out by J. G. Anderson (University of Michigan) in summer 1976 at Palestine, Texas. Both concentrations of Cl and ClO were determined by Cl atom atomic resonance fluorescence, which has been successfully used also for O atom and OH radical measurements in the stratosphere, by Anderson (*Geophys. Res. Lett.*, **2**, 231; 1975; *Geophys. Res. Lett.*, **3**, 165; 1976). ClO radicals were reacted with nitric oxide to give Cl atoms, which were detected by atomic resonance fluorescence. The technique was adapted from flow methods for laboratory measurements of Cl atoms and ClO radicals that were developed recently (see Bemand and Clyne, *J. Chem. Soc. Faraday II*, **71**, 1132; 1975). Anderson's first Cl and ClO experiments have not so far led to satisfactory results, although the prospects for the atomic resonance and other methods look promising.

Additional direct verification is in principle obtainable from measurements of ultraviolet light intensity at ground level. Whilst such measurements have been made for a number of years, the natural variation of ultraviolet levels at the ground is large. The effects of CFM would probably only become observable in this way when

the ultraviolet level was already increased by a significant, probably unacceptable amount.

Data of type (2), the rates of chemical reactions, are considered carefully in the NAS panel report. One basic difficulty was to identify correctly the major reactions, of which there are a large number. The recently-proposed role of chlorine nitrate, ClONO₂, has necessitated the introduction of several new reactions. The elementary reaction rates of chlorine nitrate are not, however, all well-defined. Fortunately, the model for ozone depletion is not very sensitive to these rates, and some confidence may be attached to the NAS panel's conclusion that the ozone depletion is reduced by a factor of about 1.7 by inclusion of chlorine nitrate in the model.

The NAS panel conclude that the largest uncertainty in the ozone depletion model is due to uncertainties in the relevant chemical reaction rate constants. They carefully identify the reactions introducing the greatest uncertainty as

$$OH + HO_2 \rightarrow H_2O + O_2$$
followed by
$$OH + HCl \rightarrow H_2O + Cl.$$
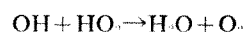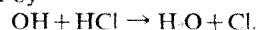
These reactions are involved in the sink processes by which active chlorine

(Cl and ClO) is removed from the ClO_x cycle. Surprisingly, the two reactions of the ClO_x cycle, that is $Cl + O_3 \rightarrow ClO + O_2$ and $O + ClO \rightarrow Cl + O_2$, do not introduce large errors, because their rates have been carefully measured in the laboratory. The first measurements of $Cl + O_3$ and $O + ClO$ rates by Bemand, Clyne and Watson at Queen Mary College in 1973 (see *J. Chem. Soc., Faraday I,* **69,** 1356; 1973; Clyne and Watson, *J. Chem. Soc. Faraday I,* **70,** 2250; 1974) have been only slightly altered as a result of subsequent work in the USA and in this laboratory.

Significant uncertainty ($\pm$ a factor of 1.7) in the prediction of ozone depletion is also introduced by the one-dimensional model used by the NAS panel to approximate the distribution and transport of the chemical species involved. This model is equivalent to averaging the concentrations, motions and chemical reactions over latitude and longitude, leaving only their dependences upon altitude and time. This approximation is considered to be a

good one. For instance, the equivalents of averaged gas transport rates appear in the model as the vertical eddy mixing or transport coefficient, which depends only on the altitude.

One further very important question concerns the effects of an increased ultraviolet intensity (UV-B, 280 to 320 nm) at ground level on animals and plants. The biological effects of UV-B are not as extensively dealt with by the NAS reports as are other aspects. This may be due to the relative lack of knowledge in this area. It is clearly to be identified as a priority area for future work on the CFM-ozone problem. However, in general terms, the deleterious effects of UV-B radiation on living cells are well known. For example, there is a clear causal connection between the more common (but less serious) forms of skin cancer (basal and squamous cell carcinomas) and UV-B exposure. The connection between the less common, but often fatal malignant melanoma and UV-B, is suggestive but not certain, according to

the NAS report.

Release of pollutants other than CFM, which can also affect ozone levels, is another topic requiring extensive future work. Among halogen compounds, further work on methyl chloride formed by the burning of vegetation, is indicated. Also, Crutzen has suggested that methyl chloroform (CH_3CCl_3), whose manufacture as a degreasing agent is considerable and possibly increasing, may eventually cause ozone depletion comparable to that of CFM 11 and 12. Bromine compounds, released for example as methyl bromide in soil fumigants, are probably not produced in large enough quantities to have major effects. However, Br atoms and BrO enter into a similar, but more effective cycle to that of Cl and ClO.

$$Br + O_3 \rightarrow BrO + O_2, \quad O + BrO \rightarrow Br + O_2.$$
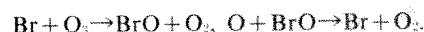
The rate constants for the bromine reactions are not particularly well known. Nitrous oxide, apparently re-

---

Cultivation of human malaria

from F. E. G. Cox

THE *in vitro* cultivation of malaria parasites has been possible for many years but only one or a few successive cycles of development have been achieved and these successes have been mainly with the malaria parasites of monkeys. The amount of physiological and biochemical knowledge that has accumulated from such studies has been vast and it cannot be disputed that much of it is directly or indirectly applicable to chemotherapeutic and immunological studies aimed at the control of malaria. What is desperately required, however, is the long term continuous cultivation of human malaria parasites which cannot be maintained in standard laboratory animals. For some time now there have been rumours of the successful cultivation of *Plasmodium falciparum* in the United States and now two independent reports have appeared, giving the long-awaited experimental details (Trager and Jensen, *Science,* **193,** 673; 1976; and Haynes *et al.,* page 767 in this issue of *Nature*). There can now be no doubt that the continuous cultivation of *P. falciparum* can be achieved in simple media and that it should shortly be possible to obtain the yields of parasites necessary for the production of experimental vaccines.

Both media are remarkably simple. Trager and Jensen use a modification of an earlier medium (RPMI 1640 with HEPES buffer and serum) for the cultivation of *P. coatneyi* from monkeys (Moore *et al., J. Am. med. Ass.,* **199,** 519; 1967) but substitute human serum for the monkey serum originally used. In essence, the medium

flows over a shallow bed of red blood cells which are disturbed as little as possible in an atmosphere of low O_2 and high CO_2. The technical details are still being modified and Trager and Jensen began by using 7% CO_2 and 5% O_2 but changed this to 1% O_2 in later experiments with better results. The technique of Haynes *et al.* is more traditional, involving a supplemented medium 199 with 1% foetal bovine serum in an atmosphere of 3% CO_2. No doubt many other ways will be found to grow *P. falciparum,* and Trager and Jensen report success using their medium in a candle jar (a container in which a candle has burned and extinguished).

The most important features of the cultivation techniques described seem to be the high levels of CO_2 and the fact that the number of red cells infected rarely rises above 1%, thus preventing the culture medium from becoming too acidic as a result of the metabolic activities of the parasites, which is a problem inherent in this kind of experiment. Although the multiplication of parasites recorded is low (a four-fold increase per cycle is assumed by Trager and Jensen) the numbers of parasites available for collection can be considerable if parallel cultures are set up and initiated from diluted cultures

at the end of each cycle. Haynes *et al.* obtained eleven successive cultures and Trager and Jensen more than twenty-five (and Trager has since reported continuous cultivation over 105 d, implying fifty or more cycles) so there is no reason why such cultures should not be maintained indefinitely.

The technique described by Trager and Jensen looks more promising than that of Haynes *et al.,* but the latter authors have made a major advance by showing that their cultures can be initiated from frozen blood samples. It should therefore be possible to initiate a culture and to freeze stabilates from that culture for the initiation of others. This would save the lives of countless owl monkeys (*Aotus trivirgatus*), which at the moment are the only sources of laboratory maintained *P. falciparum,* and also standardise experiments carried out in different laboratories.

The implications of these experiments are considerable, for as well as the biochemical and physiological observations which are now made feasible it should be possible to grow the numbers of merozoites (the stages that pass from one cell to another) needed for producing vaccines. The only homologous vaccine with real potential, in monkeys at least, is that based on merozoites (Mitchell, Butcher and Cohen, *Immunology,* **29,** 397; 1975) and the production of a suitable vaccine free from viruses is only possible from cultured parasites. With 96 million cases of malaria each year and with *P. falciparum* being the major killer the production and evaluation of such a vaccine should not be delayed.

leased by bacterial action on nitrogenous chemical fertilisers, is fairly stable in the troposphere. The possibility of this compound affecting stratospheric ozone levels should certainly be studied.

All in all, the ozone layer is fragile, and wide-ranging long-term studies of possible effects of pollutants not yet identified are essential. This should include the acquisition of more information on ozone reaction rates. Exclusive concentration on CFM 11 and 12 effects is unwise, since it is quite possible that other deleterious compounds may be present now, or at some future time, due to shifts in industrial production.

I have emphasised here the effects of CFM on the ozone layer and the consequential increase in UV-B radiation at the Earth's surface. It should also be noted that other climatic changes due to CFM cannot be excluded. For instance, Ramanathan (*Science*, **190**, 50; 1975) has stated that CFM 11 and 12 are sufficiently effective absorbers of infrared radiation (near 10 μm) to have a potentially significant effect on the Earth's heat balance.

The role of British science in the CFM-ozone problem is an important one. Much of the basic work leading to the original theories of Stolarski and Cicerone and Molina and Rowland, was carried out in the UK. Lovelock developed a sensitive electron-capture gas chromatograph, with which he was able to demonstrate the presence of various halocarbons arising from man's activities, and to obtain measurements of CFM 11 and 12 at various locations. The programme of work (COMESA) sponsored by the Department of Trade, and carried out partly in university laboratories and partly by the Meteorological Office on possible effects (slight) of Concorde on the stratosphere also

provideed valuable input data for the CFM-ozone problem. Current initiatives in the CFM-ozone problem are being taken by industry; research on both CFM 11 and 12 effects, and on substitute CFMs is being carried out. Basic science relevant to stratospheric ozone is being supported by the Science Research Council, and recently the Natural Environment Research Council has announced its intention to concentrate more resources on atmospheric science. The financially small but scientifically significant support over many years in this area by the Royal Society through its Gassiot committee has also been very valuable.

With this expertise available, it is sad therefore that there is no sign at present that the British government is planning any major programme of investigation in stratospheric science related to the CFMs and ozone, although it does contribute through the EEC organisations to the European effort. No more than £100,000 would probably be allocated to such a new investigation, if mounted eventually by the Department of the Environment. Although the restraints of the present economic situation have to be acknowledged, it is a pity, nevertheless that it was not possible to use the experience gained in the COMESA research as the basis of an ongoing and centrally-planned programme of work on atmospheric pollution.

There is a message from all this for those who ask for more "relevant" research programmes, related to current needs of society. Without the undirected, basic research programmes which were carried out, the chances are that the risks to ozone from CFMs would not have been identified until much later, if at all. Identifying the problem is a large part of finding the solution. □

sequences in pure form, by cloning in a bacterial host, has opened up the possibility of studying the sequences flanking the gene which may be important in regulation, and which cannot be obtained sufficiently pure for sequence studies by any other means. The sea urchin (*Psammechinus*) histone gene repeat unit has been cloned in phage lambda and some 25% of the nucleotide sequence determined by a group at the University of Zurich (Walter Schaffner).

The mammalian globin genes are also under determined attack in several laboratories. Charles Weissman (University of Zurich) and Richard A. Flavell (University of Amsterdam) reported the partial purification of the + and − strands respectively of rabbit globin DNA. Weissman and colleagues have isolated the + strand by hybridisation to immobilised globin cDNA; the − strand has been isolated by hybridisation with mercurated mRNA. The final purification step will be to hybridise the two strands, since only complementary sequences should anneal. This reconstituted DNA inserted into a plasmid should provide a source of the complete globin gene sequence including the flanking non-transcribed regions.

Still looking to the future, Bob Williamson (St Mary's Hospital Medical School, London) described a possible use for cloned globin gene sequences as chain-specific probes for the study of certain inherited defects such as the thalassaemias.

Recombinant DNA is also a powerful new tool for probing the immunoglobulin genes. As the first step to tracing changes in the Ig-gene coding sequences predicted by the theory of somatic mutation, Bernard Mach and colleagues (University of Geneva) have constructed plasmids carrying cDNA copies of a mouse myeloma light chain mRNA. The Geneva group has obtained plasmids containing constant and variable region sequences both separately and together. Hybridisation studies have shown that the untranslated region in kappa light chain messenger from different sources is highly conserved.

However, these experiments exemplify some of the conjectural risks which have caused concern about this type of work. The mRNA was not pure and indeed 40% of the initial recombinants contained unknown sequences. More worrying still is that myeloma cells are known to express C-type RNA viruses and that the plasmid carried a drug-resistance marker containing an IS sequence.

This problem is avoided in work from the Laboratory of Molecular Biology in Cambridge reported by Terry Rabbitts. Light chain mRNA

Nuts and bolts of genetic engineering

by Eleanor Lawrence

> The Biochemical Society's 9th Harden Conference, entitled the Role of Recombinant DNA in Molecular Biology, was held at Wye College on September 20–24, 1976.

IN the present state of the art, a meeting on *in vitro* recombinant DNA must of necessity concern itself more with means than ends. This is partly a result of the need to develop and refine a range of completely new techniques; but it is also a consequence of the almost unique situation in which the novel technology has been applied less

than whole-heartedly, as a result of the voluntary postponement of some experiments until experimental conditions minimising the hypothetical safety risks of this work had been agreed.

But now that experimental guidelines have been published in the United States and Britain, workers are preparing to clone genes from everything from maize to man. A foretaste of the power of these new techniques to investigate otherwise unapproachable problems of the organisation of the eukaryotic genome was given at the meeting.

The ability to obtain specific gene

from myeloma cells can be reverse transcribed into cDNA by starting with a hexanucleotide primer T_2G_3T that has been shown to prime only at a site just within the 3' end of the C region. In this way pure fragments containing the immunoglobulin V and C coding sequences only can be made, without the need for constructing plasmids by way of impure mRNA. A comparison of the reverse transcriptase "stutter" pattern of mRNA from a mouse T-cell lymphoma has produced evidence that in these T cells the immunoglobulin genes are at least transcribed.

The most elegant exploitation of the recombinant DNA technology was in work by David Hogness's group (Stanford University) on the organisation of the *Drosophila* chromosome. Taking advantage of the relatively small size of the *Drosophila* genome and the availability of polytene chromosomes, they have constructed plasmids containing different *Drosophila* gene sequences which have been mapped on the chromosomes. Hogness described analyses of the histone genes, the genes induced by 'heat shock', and of a sequence coding for an abundant mRNA. It is particularly intriguing that this mRNA sequence is found by *in situ* hybridisation to be located at 33 sites dispersed throughout the genome. What the dispersed gene sequence means in terms of evolution and control is still a matter for dispute. Hogness tends to the view that each individual sequence will be expressed in a different cell type, which would explain the maintenance of 33 separate sequences coding for the same product.

For those who wish to study the functional expression of eukaryotic genes, cloning in plasmids or phage has one overwhelming disadvantage. So far there are only a few reports that eukaryotic genes can be expressed in bacteria. Also, there will eventually be a need to put characterised, specfically mutated genes back into eukaryotic cells to study their expression. This might be done using animal tumour viruses that can integrate into mammalian cell chromosomes.

The potential of SV40 as an 'eukaryotic vector' is at present being explored and Paul Berg (Stanford University) described the construction of an SV40 vector in which the late genes have been deleted, and its use to clone fragments of lambda DNA in monkey cells. Additional vectors are under construction in which the early region has been deleted; these may more readily allow transcription of the inserted DNA. All these SV40 vectors are dependent on a temperature-sensitive helper virus for growth.

In this system there is however an ultimate constraint on the size of the inserted DNA because of packaging into SV40 protein coats of defined size. This problem may be overcome in a system that Joe Sambrook (Cold Spring Harbor Laboratory) is developing. He has found that human cells transformed by SV40 virus or DNA contain a large number of free SV40 DNA molecules. Transformation rather than plaque formation could be used as the selective marker for recombinants, and since no helper virus is needed and no infectious virus is released less rigorous containment conditions are required.

A system in which cloned eukaryotic DNA might be transcribed and translated was described by Janet Mertz (Laboratory of Molecular Biology, Cambridge). Both prokaryotic and eukaryotic DNA injected into the nucleus of *Xenopus* oocytes appear to be reconstituted into chromatin and transcribed but whether the transcripts are sensible or are translated is still problematical. One use for this system would be in identifying eukaryotic promoters by fragmenting well-characterised genes and then determining which pieces are transcribed.

So far most attention has been focused on the use of restriction enzymes for mapping gene sequences and producing recombinant DNAs. But there is another side to the story as pointed out by Adrian Bird (MRC Mammalian Genome Unit, Edinburgh) who has been studying the natural methylation of some restriction targets, in the ribosomal DNA of *Xenopus laevis*, which renders them resistant to attack. By analysis with a set of restriction enzymes he has shown that the probability of methylation at each suitable site is high (0.99) but that there is a unique site within the rDNA repeating unit that remains unmethylated in up to half of repeats. He has also shown that the pattern of methylation is copied onto newly synthesised DNA *in vivo* and so is inherited by a daughter somatic cell from its parent.

The prospect of using eukaryotic genes cloned in bacterial cells to make large amounts of valuable gene products is still very much in the future. But Noreen Murray (University of Edinburgh) described ways of harnessing the enormously efficient lambda leftward promoter to provide enhanced expression of inserted bacterial genes, a technique which could be applied to produce large amounts of many useful phage and bacterial enzymes.

Although the issue of containment levels and possible safety risks hardly surfaced officially, safer experimental procedures were much on everyone's mind. Paul Berg expressed concern that antibiotic resistance was still so widely used as a selective marker since

it could contribute to the persistence of any escaped vector, not only in the human intestinal tract but in the general environment (Sydney Brenner has been quoted as remarking that the world is a dilute solution of tetracycline). The proliferation of different vectors for cloning recombinant DNA also led Berg to propose a register clearly setting out their desirable and undesirable features (such as the possession of insertion sequences). Berg said that he would collate a list, from names and addresses sent to him, of people working with vectors. This would be published in *Nucleic Acid Recombinant Scientific Memoranda*, produced by and available from the National Institutes of Health. □

Actins plural

from Dennis Bray

The impersonal air that is usual in a comment like this wears thin when its author is forced to eat his words. For it was indeed I who solemnly predicted, three years ago, that the actin present in different kinds of cells was one and the same gene product (Bray, *Cold Spring Harbor Symp. quant. Biol.*, **37**, 567; 1973); and in this I could not have been more wrong. Hardly were the words out of my mouth when new forms of actin began to spring up like mushrooms. So well they grew, and in such numbers, that little remains now but to turn to Philosophy and learn the value of the refutation of a cherished hypothesis. (Popper, *The Logic of Scientific Discovery*, Hutchinson, 1959).

It has always been clear that the actin and myosin of non-muscle cells must behave in a different fashion to that of muscle: the exigencies of motility and cell division ensure that. But evidence for differences in amino-acid sequence, rather than in the modifying influence of ancillary components, is harder to come by. Some early hints exist in the literature, such as that in Fig. 12 of a paper on the actomyosin from brain (Berl and Puszkin, *Biochemistry*, **9**, 2058; 1970). This shows that brain actin and muscle actin run differently on urea-acrylamide gels and seems to have been overlooked in recent papers that present the same fact (Storti and Rich, 1976, *Proc. natn. Acad. Sci. U.S.A.*, **73**, 2321; 1976; Storti, Coen and Rich, *Cell*, **8**, 521; 1976). But it was probably in the course of a thorough examination of platelet actin, that the first unambiguous differences—in the form of cyanogen bromide peptides—were found (Booyse, Hoveke, and Rafelson, *J. biol. Chem.*, **248**, 4083; 1973). A later comparison of peptides from brain actin to those from muscle also showed differences (Gruenstein and Rich, *Biochem.*

biophys. Res. Commun., **64**, 472; 1975).

At this point it was still possible, with difficulty, to believe in one actin. The cyanogen bromide evidence was clear, but perhaps there had been an incomplete cleavage. Criticism of the work on brain actin was possible, since the data showed as much variation between different samples of muscle actin as between muscle and brain. But there was little point in chopping these actins down: no sooner the attempt than two stood where previously there had been but one! (Dukas, *The Sorcerer's Apprentice*, 1897). Evidence was adduced, first for platelets and then for other cells, that not only was their actin distinct, but it was multiple. The platelet actin that Booyse and his collaborators studied was extracted as if it was muscle actin, and it obligingly behaved in this way. But this was only a small fraction of the total actin and when a different approach was tried, a protein with a quite different ability to polymerise was obtained. (Abramowitz, Stracher, and Detwiler, *Arch. Biochem. Biophys.*, **167**, 230; 1975). Moreover amoeba, fibroblasts and brain, also give small amounts of polymerisable actin, or large amounts of unpolymerisable actin, according to the extraction used. There is still a lurking question of whether a slight denaturation is involved, but on the face of it this is evidence for two actins.

An even greater virtuosity is apparent in developing muscle which shows no less than three actins (Whalen, Butler-Browne, and Gros, *Proc. natn. Acad. Sci. U.S.A.*, **73**, 2018; 1976). The evidence here is from gels lacking SDS—and it must be recalled that in some conditions even pure muscle actin wil split into a number of bands (Pardee and Bamburg, *J. Neurochem.*, **26**, 1093; 1963; Carsten and Mommaerts, *Biochemistry*, **2**, 28; 1963)—but the changes that accompany differentiation are convincing. Finally, lest there are any lingering doubts about the reality of the differences considered, comes definitive evidence from the amino acid sequencers: slow, but exceedingly hard to demolish. With the complete sequence of muscle actin in hand, Elzinga and colleagues have now compared peptides from human heart and human platelet actin (Elzinga, Maron, and Adelstein, *Science*, **191**, 94; 1976). A diminutive difference exists—1 residue in 29 is changed, from a harmless valine to an innocuous threonine—but it is enough.

From the wider biological standpoint these results add to other examples in which closely related proteins are used to perform similar, but distinct, functions. The evidence available for tubulin suggests that the various organelles in which this protein appears are based on different gene products;

and a similar statement probably holds for myosin and tropomyosin. Undoubtedly there are advantages in this: the process of gene duplication is familiar in haemoglobin, and provides a ready-made protein on which evolution can go to work. But, if Cassandra is allowed one parting shot, in the matter of the disposition of these proteins within the cell many mysteries remain. It seems scarcely credible that such similar molecules could be moved within the same small space and yet not cross-react or finish in the wrong place.

Southampton model fair

from N. MacDonald

A research conference on nonlinear mathematical modelling was held at Southampton University on September 13–17, 1976. It was sponsored by the Social Science Research Council and organised by D. R. J. Chillingworth, H. B. Griffiths and D. A. Rand of the mathematics department at Southampton.

FROM the viewpoint of a physicist working on non-linear modelling in biology, the most satisfying talks dealt with specific problems from the hard sciences, handled with mathematical sophistication. There has been much interest recently in the behaviour of travelling waves in solutions undergoing oscillatory chemical reactions, such as the Zhabotinskii reaction, both as objects of study in themselves and as a possible paradigm for pattern formation in organisms. N. Kopell (Northeastern University, Boston) described two elegant applications of bifurcation theory, to the outward propagation of chemical waves from a centre, and to the collision of two such waves.

A considerably deeper physical problem was the subject of a fascinating lecture by G. Jona-Lasinio (University of Rome). He described recent work by the Russian mathematicians Bleher and Sinai on phase transitions (*Commun. Math. Phys.*, **45**, 247; 1976; see also a review by Jona-Lasinio in *Nuovo Cimento*, **26B**, 99; 1976). These are difficult phenomena to model because one has to build up from local assumptions about interactions between atoms to an essentially collective large scale switch in the behaviour of the material. Bleher and Sinai treat this scaling process as an iterative operation on the probability distribution. The asymptotic large scale forms of the probability distribution are obtained as the fixed points of a discrete dynamical system on a function space, and phase

transitions are related to the bifurcations of these fixed points.

The bifurcation theory of discrete dynamical systems is a very popular subject at present. The great richness of the behaviour of solutions of even the simplest difference equations (discrete dynamical systems on the real numbers) was the subject of a recent review by R. M. May (*Nature*, **261**, 459; 1976). At this meeting J. Guckenheimer (University of Santa Cruz) described the application of this kind of mathematics to a model of the population fluctuations in the famous blowfly experiment of Nicholson.

Turning to the more qualitative surveys, B. Goodwin (University of Sussex) discussed morphogenesis, in particular recent results on the circumferential field in the development of the insect limb. He stressed the gap between the field concept, which is effective in coordinating data, and the underlying mechanisms, as well as that between the abstract dynamics suggested by catastrophe theory and the observable expression of form in the organism. Evidence that the conceptual difficulties involved in understanding complex biological systems are of a different order from the mathematical difficulties of current modelling techniques, came in the lecture by R. Rosen (Dalhousie University).

Several of the lectures on the social sciences had some common themes. A. C. Renfrew (Southampton University) dealt with attempts to set up a framework in which questions about the growth and decline of cultures can be handled by dynamical modelling techniques, analogous to those of the Forrester–Meadows group. K. L. Cooke (Pomona University, California) described the application of one such method to Renfrew's description of the Aegean culture. A. G. Wilson (University of Leeds) gave an account of the great variety of modelling techniques that have been applied to urban geography, and stressed the need to move on to a more integrated approach.

P. J. Harrison (University of Warwick) discussed the problem of short term prediction in economic and social contexts. Here standard methods apply essentially to linear behaviour, while the need to handle non-linearity becomes apparent especially when concerned with the need to keep options open for the possible reversal of an existing policy.

Although the main aim of the meeting was to set up an interaction between pure mathematicians and modellers, it was evident that cross-disciplinary contacts between modellers were equally stimulating. These were mainly informal, although some structure was given by a series of meetings in the archaeology department. All

open-minded members of the conference must have taken away, as well as their copies of Poston and Stewart's little red book, and the good resolution to think more geometrically, much benefit from the chance to talk with people from a remarkably wide range of disciplines. □

Venus is another differentiated planet

from G. Fielder

MANY scientists were surprised to find that the Moon was a differentiated planet. Now, *in situ* measurements of parts of the surface layers of Venus have indicated that the Earth's "sister" planet is more like the Earth than the Moon in its radioactive element content. Whereas measurements of the potassium, uranium, and thorium concentrations made in October 1975 using Veneras 9 and 10 indicated the presence of basaltic type rocks, earlier results from Venera 8 (July 1972) had indicated the presence elsewhere of rocks similar in their radioactivity to that of a typical terrestrial granite.

Writing with K. P. Florenskii, A. T. Bazilevskii and A. S. Selivanov in *Doklady Akad. Nauk. SSSR* (**228**, 570; 1976) the late A. N. Vinogradov has reported studies of the photographic panoramas and other data transmitted from Veneras 9 and 10 by way of a relay satellite. Telecameras mounted 1 m above the surface took high Sun photographs, including the one reconstructed in Fig. 1, with a resolution of about 1/3 degree. So effective was the

dense atmosphere of Venus in diffusing the weak light at ground level that high contrast versions of the photographs show "artificial shadows" which always point towards the appropriate camera.

The upper right hand corner of Fig. 1 shows the skyline, estimated to be several tens of metres distant. The sharpness of the skyline indicates a fairly high atmospheric transparency and, consequently, a comparatively low dust content. If any wind had been blowing across the surface when the pictorial data were transmitted, the wind had not been dust laden. Nevertheless, the authors have suggested that the clod-like nature of the dark soil filling the spaces between the numerous rock blocks may have resulted, in part, from the evacuation of fine particles by a weak wind. An alternative explanation, they suggest, is that the clods were produced following hydration of the particles.

Typically, the rock blocks themselves have dimensions of some tens of centimetres. Some are angular, others rounded. Many indicate that they have been detached from hard, layered bedrock. Others, which appear to be mostly buried, could even be outcrops of bedrock. A few rocks have cracks which cross the bedding plane. Others appear to be vesicular.

Together with the knowledge from radio observations that the surface layers of Venus are underdense, the picture data establish that certain dynamic processes are in operation there now. If the corrosive agents hydrochloric acid and hydrofluoric acid observed in the upper atmosphere of Venus are present at ground level they could be an important factor in the rounding and disintegration of the

rocks to generate the soil. Whatever the predominant mechanism of denudation on Venus, it is clear from the abundance of sharply sculptured rocks on the surface where Venera 9 landed that this is a geologically young area of the planet.

The clouds of Venus, which prevent our viewing its surface directly, are found at up to 70 km above the surface. Most meteoroids of the present day populations would probably disintegrate in the thick, dense atmosphere (97% CO_2) before reaching the surface so that recent disturbances of the rocks of Venus are likely to be of internal origin. Now that two panoramas of the surface of Venus have shown scenes not unlike those of terrestrial desert landscapes, the next step will be to discover the processes which moulded the landforms of Venus. □

Fig. 1 Part of Venera 9 panorama. The rotation axis of the telecamera was inclined at ~40° to the vertical axis of the landing stage. The rocks shown here lie on, and form, a steep slope (the on-board inclinometer reading was ~30°). Data auxiliary to the picture data were transmitted in vertical bands which have been removed.

So Madagascar was to the north

from Peter J. Smith

IN a recent article on the Madagascar controversy, Kent and Tarling (*Nature*, **261**, 304; 1976) were adamant in claiming that "the Madagascar controversy still lives". They were right of course, even excluding the element of self-fulfilling prophesy. To expect everyone to accept that "palaeomagnetism has provided the definitive answer" to the Madagascar problem (*Nature*, **259**, 80; 1976) was evidently overoptimistic, or at least premature. The controversy does indeed live. But for how much longer in the light of a new palaeomagnetic analysis from McElhinney *et al.* (*Geology*, **4**, 455; 1976)?

The long-standing argument is, of course, over the geographical position of Madagascar before the breakup of Gondwanaland. Did this large island lie in the coastal embayment of Mozambique some 4° south of its present position? Or was it about 15° north, adjacent to the coast of Kenya-Tanzania-Somalia? Or has it always been where it is with respect to Africa, except perhaps for just a small amount of eastward drift? The answer is critical in defining the fit of eastern and western Gondwanaland. Evidence of various types has been adduced in favour of all three positions at various times, although clearly no more than one of the options can be correct.

The problem seemed well on the way to solution earlier this year when Embleton and McElhinney (*Earth planet. Sci. Lett.*, **27**, 329; 1976) presented palaeomagnetic data supporting the southward drift of Madagascar—

that is, a predrift position to the north. But Kent and Tarling were not so sure, citing a major geological argument against a northern position. In the vicinity of Kenya-Tanzania there are thick sediments apparently deposited from the Upper Carboniferous onwards, both onshore and out to several hundred kilometres offshore. So how, Kent and Tarling asked, could these deposits "exist in their present position if this was occupied by a microcontinent that has since moved away"?

They then went on to question both the validity of the palaeomagnetic data and the significance of the analysis based on them. The point about validity hinged on whether or not the measured sediments had been remagnetised since they were laid down. But Kent and Tarling argued that even if remagnetisation could be ruled out and the data were perfectly valid, there is at least one alternative way of comparing the Madagascar data with those from other landmasses. And this other way suggests that a southward drift of Madagascar is no more likely than a northward drift or no drift at all. The point was not that one method is any better or worse than the other, but simply that both are equally valid and reasonable.

Be that as it may, there is now a sequel. Since Embleton and McElhinny reported their first data, palaeomagnetic investigations of Madagascan rocks have continued (both in Australia and France) to the point where there is now an almost complete palaeomagnetic coverage of the Karroo Supergroup from the Upper Carboniferous to the Middle Jurassic. And all of the results have been brought together and analysed by McElhinny *et al.*

The Madagascan data fall into three sets covering, respectively, the time-spans: (1) Upper Carboniferous to Lower Permian, (2) Upper Permian to Lower Triassic and (3) Middle Triassic to Middle Jurassic. Unfortunately, appropriate African data are heavily biassed towards the Middle Triassic-Middle Jurassic, which makes comparison over the total Karroo period difficult.

And the results of the comparison are striking. Taking the time-spans (1), (2) and (3) together, the average angular difference between the Madagascan and African–South American pole positions is 16.8° for Madagascar in the southern predrift position and 15.3° assuming the island to have maintained its present position. These are palaeomagnetically significant differences. By contrast, the average angular difference for Madagascar in the northern predrift position is only 4.5°, which is palaeomagnetically insignificant. Moreover, these averages conceal

no anomalies; time-spans (1), (2) and (3) taken separately each give a similar result.

In short, the palaeomagnetic case for the northern predrift position of Madagascar is, in the word of McElhinny and his colleagues, "unequivocal". Even the sceptics must surely agree that at the very least the case is strong. It is difficult to believe that the palaeomagnetic data are not valid. Remagnetisation, for example, is possible, but hardly likely in view of the similar results for the three time-spans and the number of different rock types involved. So it would appear that someone has a lot of reconciling to do. □

Competition and sexual dimorphism in plants

from Peter D. Moore

As members of a species in which males and females are increasingly coming into competition with one another, it is interesting for us to consider the evolutionary consequences which have resulted from such a state of affairs in other species. Sexual dimorphism is the term used for situations in which the male and female of a species are adapted so that they draw on different resources, or tap a resource in different ways. It results from conditions where the limitations of the environment can produce deleterious competitive interactions between males and females inhabiting the same area.

Sexual dimorphism is quite common among animals and may result in considerable differences between the sexes. In many birds of prey, such as the sparrow hawk (*Accipiter nisus*) the male is considerably smaller than the female. The two sexes feed on different prey species, the male concentrating on smaller, more agile species, and the female on larger ones.

Among plants, very few examples of sexual dimorphism have been described. Obviously it could exist only within dioecious species, where the sexes are separate. One example is the sheep's sorrel (*Rumex acetosella*) which was studied by Putwain and Harper (*J. Ecol.*, **60**, 113; 1972). Here the two sexes were found growing together, but showed considerable differences in their times of growth and in the allocation of food reserves. Male plants grew faster in the spring, but expended less energy on the production of inflorescences and pollen than did the female on seed production. Female plants were later in their growth, but finally developed a taller canopy. Males

invested more energy in the production of roots and vegetative offshoots. Males and females thus tap the environmental resources at different times and in rather different ways, constituting an example of niche diversification between the sexes.

Freeman, Klikoff and Harper (*Science*, **193**, 597; 1976) have examined several dioecious plant species in relation to soil moisture in arid to semiarid habitats in Utah. They found that the ratio of male : female plants in all species examined varied with environmental conditions, mainly water availability. For example the meadow rue (*Thalictrum fendleri*) had a male: female ratio of almost 7 : 1 in dry sunny positions, whereas in moist shady conditions it was nearer 1 : 4. Mormon tea (*Ephedra viridis*) had a male : female ratio of 1.6 : 1 on steep, arid slopes and about 1 : 2 on alluvial bottomlands. The salt grass (*Distichlis spicata*) gave a ratio of about 3 : 1 in very saline conditions and about 1 : 1.5 in moderately saline areas.

These data suggest that males of all these species survive better than females in conditions of water stress, whereas the reverse is true where water is more available. Alternatively, sex determination may be environmentally controlled. Whatever the cause of this state of affairs, Freeman *et al.* suggest that the differentiation of niches between the sexes has adaptive significance. Reproductive success in males in these wind-pollinated species may be enhanced by their being located on exposed, dry ridges. Seed output from females, on the other hand, would be greater in conditions of improved water supply, especially later in the season.

They have not examined the role of vegetative reproduction in determining sex ratios. Putwain and Harper showed the importance of this influence in their *Rumex* studies and it is possible that resource allocation to vegetative reproduction in these species varies between the sexes and is differentially affected by water availability. *Ephedra*, for example, spreads extensively by vegetative reproduction. Niche differentiation between the sexes of these plants has evidently occurred in response to a selective pressure, but the precise physiological or morphological cause of the variations in sex ratio remain unclear.

Dioecious plant species are not uncommon and it will be interesting to see whether more of them exhibit sexual dimorphism. Sex ratios in populations, however, may not always be the best guide to dimorphism in plants because of the influence that vegetative reproduction can have by redressing the effects of differential germination and survival. □

review article

Transposable genetic elements and plasmid evolution

Stanley N. Cohen*

Transposable elements of DNA that are structurally defined and genetically discrete units seem to have an important role in the evolution of bacterial plasmids. Recombination occurring at the termini of such elements can result in the joining together of unrelated DNA segments that lack extensive nucleotide sequence homology. In addition, transposable elements serve as novel biological switches capable of turning on and off the expression of nearby genes as a consequence of their insertion into or excision from plasmid genomes.

ALTHOUGH the existence of plasmids in bacteria has been known for more than two decades, the extraordinary prevalence and biological diversity of these extrachromosomal genetic elements has been appreciated only recently. Plasmids were first identified in the Enterobacteriaceae, but subsequently have been found in almost every bacterial group where they have been searched for. Naturally-occurring plasmids vary in size from the 2,250-nucleotide-long minicircular "cryptic" plasmid of *Escherichia coli* strain 15 (ref. 1) to the large and complex F′ plasmids[2], which may contain more than 400,000 nucleotide pairs and carry up to 600 genes. A wide variety of biological functions is specified by plasmids (Table 1).

The essential feature common to all plasmids is a replication region capable of propagating itself and any genes linked to it. Presumably, early steps in plasmid evolution involve the assembly of deoxyribonucleotides to form a site for initiation of DNA replication (the replication origin) and acquisition of genes that specify functions required for autonomous replication. Subsequent linkage of the resulting replication segment to one or more genes that provide a biological advantage to cells carrying the plasmid would be expected to aid its propagation. Later steps in the evolution of some plasmids might involve the addition of groups of genes that facilitate inter-bacterial transfer or an extracellular existence. Thus, in this evolutionary continuum, plasmids can be viewed as primitive bacteriophages that have not yet acquired those specialised functions necessary for a complex replicative cycle or the production of infectious particles capable of existing outside the host cell. Conversely, bacteriophages can be viewed as plasmids in which a basic replication region has become linked to gene segments that specify the additional biological functions needed for the production of phage particles.

Until recently, it has been generally assumed that the structural evolution of extrachromosomal elements has occurred by mechanisms of "general" or "ordinary" recombination. In *E. coli*, general recombination involves the bacterial *rec*A gene product, and is commonly dependent on physical breakage and reciprocal exchange of DNA sequences in a region of extensive genetic homology[3]. While

this mechanism has undoubtedly played some part in plasmid evolution, experiments carried out during the past several years have shown that the acquisition of certain new segments of genetic material by plasmid and bacteriophage genomes can occur by "illegitimate"[4,5] recombination of DNA sequences that have little or no ancestral relationship[6,7]. Thus, the remarkable structural and genetic diversity of plasmids seems in part to be the consequence of quantal, rather than reciprocal, recombination that involves the joining of termini of separately evolving and structurally defined segments of DNA and the translocation (transposition) of genetic segments from one DNA site to another. Moreover, at least some of the transposable elements involved in the structural evolution of plasmids and other prokaryotic chromosomes can serve also as biological switches capable of turning groups of genes on and off as a consequence of their insertion and excision. Thus, transposable elements seem to have an important role in both the structural organisation and the functional expression of genetic information on DNA.

Segmental composition of large plasmids

Investigations from several laboratories[8-12] have shown that certain large bacterial plasmids can undergo reversible dissociation into separate units in some bacterial species. For example, "cointegrate"[13] antibiotic resistance plasmids can dissociate in *Proteus mirabilis* into separate segments that carry genes for transfer functions (the RTF replicon) or for resistance to antimicrobial agents (the R determinant segment)[8-12]. Each of the two component regions of cointegrate plasmids contains its own replication origin[13]. Since independently formed isolates of an RTF unit derived from the same plasmid seem to be genetically and structurally identical[15,16], dissociation and reassociation (that is, recombination) of cointegrate plasmids must necessarily involve specific DNA loci. Electron microscope heteroduplex studies of the isolated RTF segment of one antibiotic resistance plasmid has shown that it has considerable DNA nucleotide sequence homology with the RTF regions of several other cointegrate plasmids[17]. Moreover, the RTF component of the antibiotic resistance plasmid was found to share much homology with the corresponding segments of the F and ColV-K94 plasmids.

Further heteroduplex mapping of several phenotypically

Table 1 Some properties coded by naturally occurring plasmids

Property	Exemplified by
Fertility—ability to transfer genetic material by conjugation	F, R1, Coll
Production of bacteriocins	CloDF13 (*Enterobacterium cloacae*) ColE1
Antibiotic production	SCP1 plasmid of *Streptomyces coelicolor*
Heavy metal resistance (Cd^{2+}, Hg^{2+})	pI258 (*S. aureus*), R6
Ultraviolet resistance	CollB, R46
Enterotoxin	Ent
Virulence factors, haemolysin K88 antigen	ColV, Hly
Metabolism of camphor, octane, and so on	Cam, Oct (*Pseudomonas*)
Tumorigenicity in plants	TI-plasmid of *Agrobacterium tumefaciens*
Restriction/modification	Production of *Eco*R1 endonuclease and methylase by plasmid of RY13

Plasmids listed are indigenous to *E. coli* unless otherwise indicated.

distinct plasmids has indicated that their regions of homology begin and terminate at very precise boundaries[17]: segments to one side of the boundary show major differences in DNA nucleotide sequences, whereas segments to the other side are preserved largely intact. R-determinant segments of certain antibiotic resistance plasmids isolated in widely spaced parts of the world were found to be preserved almost intact, but no detectable homology was seen with the corresponding segments of the F and ColV-K94 plasmids. Thus, the different phenotypic properties of these large plasmids seem to result in part from the linkage

Fig. 1 Inverted repeats. *a*, Electron photomicrograph of formamide-spread[81] "mini-plasmid" derivative of the pSC105 plasmid[80]. The inverted-repeat units forming the stalk (IR) of the Km-resistance hairpin-loop structure are indicated. *b*, Schematic representation of formation of hairpin-loop structure showing palindromic stem. The DNA nucleotide sequences *ABC* are repeated in inverted orientation on this DNA. Following DNA denaturation and strand separation, hairpin-loop structures can be formed by intrastrand annealing of the complementary nucleotides. The sequence bracketing the base of the stem forms an axis of twofold rotational symmetry.

a

b

```
Double   5'——A B C——X Y Z——C'B'A'——3'
strand
DNA      3'——A' B' C'——X' Y' Z'——C B A——5'
                        │ Denaturation and
                        ↓ intra-strand annealing
```

of separately derived supplemental gene segments to the basic RTF unit. These observations suggest that different combinations of separately evolving DNA segments that have little or no ancestral relationship may be involved in plasmid formation.

Inverted repeats on plasmid DNA

Heteroduplex studies of plasmid DNA sequence relationships by Sharp *et al.*[17,18] also resulted in the discovery of inverted repeats on plasmid DNA. Inverted repeats consist of DNA segments that are duplicated in reverse nucleotide sequence orientation along the genome (Fig. 1). Because of the antiparallel nature of DNA polynucleotide chains, the nucleotide sequences contained within inverted repeats are identical on each of the DNA strands from the 3' to 5' directions; consequently, a structurally defined DNA sequence and its complement exist on each strand of the DNA molecule. During heteroduplex analysis, intrastrand annealing of the primary and repeated complementary sequences results in the formation of hairpin-loop or foldback structures. The complementary sequences are contained in the stalk of such structures, whereas intervening non-duplicated sequences form the loop part. An axis of twofold rotational symmetry bisects the palindromic stem that is formed by intrastrand annealing of the inverted repeat sequences during formamide spreading of DNA for heteroduplex analysis. In at least two instances, antibiotic resistance genes carried by plasmids (that is, the genes specifying resistance to tetracycline and kanamycin-neomycin) were shown to map in the immediate vicinity of inverted repeat structures[17].

Insertion sequence (IS) regions on plasmids

One of the hairpin-loop structures located on the R6 and R6-5 plasmids was found to have another structural feature of special interest. R6-5 is a spontaneously isolated variant of R6; although it does not express the tetracycline (Tc) resistance of the parent plasmid, R6-5 was nevertheless found to contain the hairpin-loop structure associated with the Tc-resistance gene of R6[17]. Electron microscope heteroduplex analysis of the two plasmids indicated that the Tc-sensitive R6-5 plasmid contains all of the DNA sequences present in the Tc-resistant R6 plasmid—plus an additional segment of DNA 1.4 kb in length[17]. Spontaneous excision of the segment from the R6-5 plasmid, which occurs at low frequencies during normal growth of bacteria carrying the plasmid and at higher frequency[19] after transformation[20] of *E. coli* by R6-5 plasmid DNA, leads to regeneration of a plasmid that is phenotypically and structurally identical to the tetracycline-resistant R6 parent. These results indicate that a Tc-resistance gene is present on the R6-5 plasmid, but that it is being inactivated reversibly by insertion of an additional segment of DNA[17].

The inserted DNA segment that turns on and off the expression of tetracycline resistance in the R6-5 plasmid was shown to be a direct duplication of a structurally defined nucleotide sequence that is repeated again in

R6-5
(Tc-sensitive)

R6 OR S R6-5Tc^R
(Tc-resistant)

Fig. 2 Turn on and turn off of Tc-resistance gene of R6/R6-5 plasmid by excision and insertion of repeated DNA sequence. The nucleotide sequences ABC and $C'B'A'$ form a hairpin-loop structure by intrastrand annealing during spreading of plasmid DNA in the presence of formamide. Insertion of the sequence $A_2B_2C_2$, which is a direct repeat of $A_1B_1C_1$ results in loss of expression of the nearby Tc-resistance gene. Excision of this sequence leads to expression of Tc resistance by the plasmid.

reverse orientation further along the plasmid genome to form the stalk of the hairpin-loop structure near the Tc-resistance gene (Fig. 2). Insertion and excision of this segment of DNA at a site within the loop affects expression of the nearby Tc-resistance gene.

Inactivation of genes as a consequence of insertion of DNA sequences of defined length has been recognised for about a decade. Insertion sequence (IS) elements (for review, see ref. 7), which range in length from 800 to 1,400 base pairs, have been known to become integrated at multiple sites in bacterial and phage genomes and to produce polar mutations[21-25]; insertion of an IS element not only abolishes the function of the gene into which the element is inserted, but it can also affect functioning of genes distal to the insertion site with respect to the promoter of the operon. Moreover, at least one IS element carries a promoter region, and is capable of turning on or off the activity of adjacent genes according to the direction of insertion of the IS segment[25]. The 1.4-kb DNA segment observed to affect expression of tetracycline resistance in the R6/R6-5 plasmid as a consequence of its insertion or

excision seemed to be similar in its characteristics to the IS elements[17].

Subsequent electron microscope heteroduplex studies[26] have shown that the DNA sequence that forms the stalk of the Tc-resistance hairpin-loop structure of the R6 plasmid, and which is also the inserted DNA segment affecting expression of Tc resistance, is one of the previously identified IS regions—IS3. These studies have involved the use of previously characterised[27,28] λ phage mutants carrying the known IS regions; separated strands of phage λ DNA containing IS1, IS2 or IS3 were annealed with R6-5 plasmid DNA, and homology of the IS3 element with the sequence of the Tc-resistance stalk was shown[26].

Parallel investigations by Hu et al.[19] have shown that IS1 and IS2 elements are also present on R6 and R6-5 plasmids, and on the related plasmid R100. IS2 exists as a simple insertion within the RTF unit. Two copies of IS1 exist as direct repeats at the boundaries of the RTF and R-determinants of these plasmids, and by analogy from the data by Sharp et al.[17], exist also at corresponding locations of certain other cointegrate antibiotic resistance plasmids that are homologous with R6 in these regions. Using similar methods, Ptashne and Cohen[26] also demonstrated the presence of directly repeated IS1 insertions at the junction of the RTF and R-determinant units of R6-5.

The discovery of directly repeated copies of the IS1 element at the two junctions of the RTF and R-determinant components of cointegrate antibiotic resistance plasmids provides a possible mechanism for the previously reported reversible dissociation of such plasmids into separate segments[8-11]. The IS1 elements also provide a possible mechanism for amplification of the R-determinant segment and the formation of plasmids consisting of polygenic R-determinant units. As proposed by Hu et al.[29] and by Ptashne and Cohen[26], multiple IS1 copies of and R-determinant units could accumulate on cointegrate R-plasmids by recombination involving the IS1 element. According to this model, the IS1 elements could also mediate dissociation of the component of the plasmid: the number of R-determinant copies included on each of the component replicons would depend on which IS1 regions participate in the recombinational event. Recombination between the most distant pair of R-determinant units as

Cointegrate
R-factor
R-determinant component

R-determinant replicon

RTF replicon

RTF component
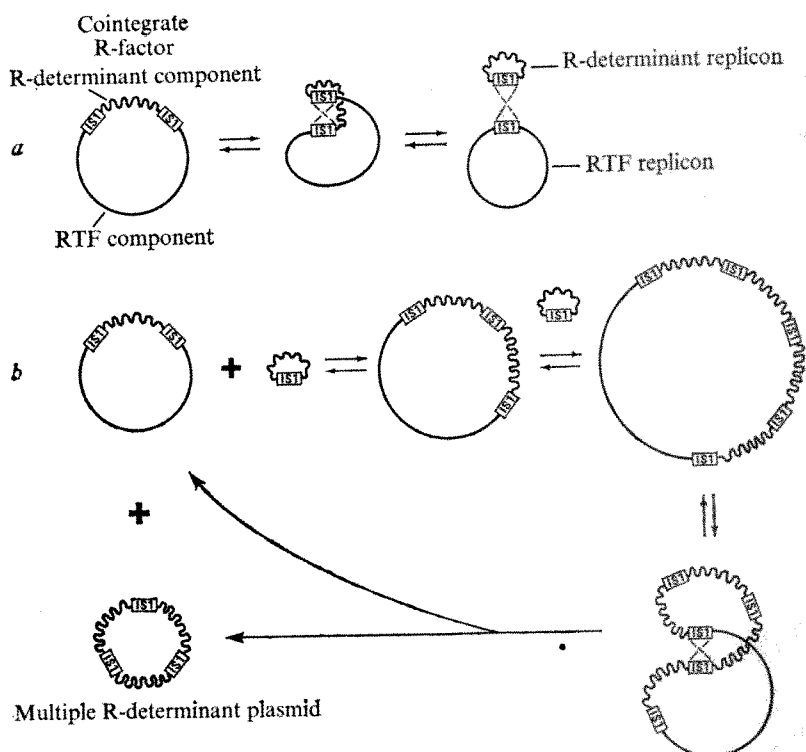
Fig. 3 Proposed mechanism[26,29] for reversible dissociation of cointegrate R plasmids at sites of IS1 elements. From Ptashne and Cohen (ref. 26). The model depicted provides for the formation of independent RTF and R-determinant replicons (a) and for the production of polygenic R-determinant units. (b). The model is schematic; it has not been determined whether both the RTF and R-determinant units receive IS elements during dissociation of cointegrate R plasmids or whether the two IS1 elements are distributed asymmetrically.

Multiple R-determinant plasmid

shown in Fig. 3 would yield an RTF replicon plus a poly-R-determinant plasmid, whereas recombination between intermediately located IS units would yield a DNA molecule containing both an RTF segment and a varying number of R-determinant copies[26]. Thus, separate RTF and R-determinant replicons[8-11], as well as plasmids containing multiple R-determinant copies[30] could be produced from cointegrate R plasmids by the same recombinational process.

IS elements also seem to be involved in the recombination of F plasmids with chromosal genes to yield F′ plasmids[31-33]. The electron microscope heteroduplex studies of Hu and Davidson and their associates[31-33] have identified three structurally defined DNA segments of F that are hotspots for illegitimate recombination events that do not require bacterial *rec*A gene function. One of these, the εζ sequence, is homologous with IS2 (ref. 29). A second recombinationally active segment on F, the αβ sequence, is indistinguishable from the IS3 element[33]. Formation of an Hfr cell by integration of F into the chromosome can occur by interaction of the plasmid with an IS element inserted in either of two polarities. In 23 of 25 independently derived Hfrs, the point of origin and the direction of transfer of the chromosome by the integrated F plasmid could be predicted by the polarity of the αβ or εζ sequence on the plasmid[32,33]. Taken together, these various data strongly suggest that IS sequences are involved in Hfr formation. These studies[31-33] suggest also that IS elements have a prominent role in the acquisition of chromosomal genes by plasmids and indicate that deletions of F-plasmid DNA can occur during F′-plasmid formation, and that one endpoint of the delegated DNA segment often is coincident with the endpoint of an IS element.

Translocation (transposition) of DNA segments carrying antibiotic resistance genes

In 1964 Dubnau and Stocker[34] reported that antibiotic resistance genes from a plasmid could become associated with the P22 prophage and subsequently attach to the chromosome of *Salmonella typhimurium*. Kondo and Mitsuhashi observed that *E. coli* phage P1 could similarly receive a resistance gene from a plasmid[35]. Later, Mitsuhashi and coworkers[37,38] described the spontaneous integration of R-plasmid genes for choramphenicol (Cm) resistance at various sites of the *E. coli* chromosome, and the subsequent movement of the Cm resistance trait on to another R plasmid or on to bacteriophage P1, which could then transfer the resistance interbacterially. Other observations from the same laboratory suggested that Tc-resistance genes

---

Abbreviated Rules of Nomenclature for transposable elements of DNA*

Definition
Transposable elements are DNA segments which can insert into several sites in a genome.

Classes of element
(1) IS (simple insertion sequences): contain no known genes unrelated to insertion function, generally shorter than 2 kb. Symbols: IS*1*, IS*2*, IS*3* and so on.
(2) Tn (more complex, transposable elements, often containing IS elements): behave formally like IS elements, but contain additional genes unrelated to insertion function, generally larger than 2 kb. Symbols: Tn*1*, Tn*2*, Tn*3* and so on. A different number is assigned to each independent isolate from nature, even if it is apparently identical to some previous isolate. The designations assigned in the accompanying table to some known Tn elements should be used in all publications.
(3) Episomes (complex, self-replicating elements often containing IS and Tn elements).

Organisms and genomes with inserted elements
(1) Organisms. When an inserting element has been introduced into a previously described bacterial strain, the new strain should be given an isolation number. Its genotype can be denoted by the genotype of the parent strain, followed by the name of the element in parentheses.
(2) Genomes. If location is known, specify gene or region in which element is inverted, followed by a number designating the particular insertion mutation, then by double colon and finally by name of inserted element; for example, *gal*T-236 : : IS*1* (IS*1* within gene *gal*T).

Central registry
To avoid duplications of numbers, all new IS and Tn

elements should be checked with a central registry before numerals are assigned to them in publications. A registry for Tn elements will be maintained by Esther Lederberg, Department of Medical Microbiology, Stanford University Medical School, Stanford, California 94305 (USA) as part of the Plasmid Reference Center.

		Designation of Tn elements*		
Element†	Plasmid origin‡	Resistance markers§	References	
Tn*1*	RP4	Ap	6	
Tn*2*	RSF1030	Ap	7	
Tn*3*	R1	Ap	8	
Tn*4*	R1	ApSmSu	8	
Tn*5*	JR67	Km	2	
Tn*6*	JR72	Km	2	
Tn*7*	R483	TpSm	1	
Tn*9*	pSM14	Cm	4	
Tn*10*	R100	Tc	3, 5	

†The symbol Tn*8* has not yet been assigned.
‡Natural plasmid from which element originated.
§Ap — ampicillin; Sm — streptomycin; Su — sulphonamide; Km — kanamycin; Tp — trimethoprim; Tc — tetracycline.

1 Barth, P. T., Datta, N., Hedges, R. W., and Grintner, N. J., *J. Bact.*, **125**, 800–810 (1976).
2 Berg, D. E., Davies, J., Allet, B., and Rochaix, J., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 3628–3632 (1975).
3 Foster, T. J., Howe, T. G. B., and Richmond, K. M. V., *J. Bact.*, **124**, 1153–1158 (1975).
4 Gottesman, M., and Rosner, J. L., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 5041–5045 (1975).
5 Kleckner, N., Chan, R., Tye, B., and Botstein, D., *J. molec. Biol.*, **97**, 561–575 (1975).
6 Hedges, R. W., and Jacob, A. E., *Molec. gen. Genet.*, **132**, 31–40 (1974).
7 Heffron, F., Sublett, R., Hedges, R. W., Jacob, A., and Falkow, S., *J. Bact.*, **122**, 250–276 (1975).
8 Kopecko, D. J., and Cohen, S. N., *Proc. natn. Acad. Sci. U.S.A.*, **72**, 1373–1377 (1975).

*Prepared by a committee consisting of A. Campbell (chairman), D. Berg, D. Botstein, R. Novick and P. Starlinger. The complete report will appear in *DNA Insertion Elements, Plasmids and Episomes* (Cold Spring Harbor Laboratories, Cold Spring Harbor, in the press).

were also capable of moving from plasmid to plasmid or from plasmid to chromosome[39-41].

In 1968, Anderson *et al.*[42] reported observations suggesting that an Ap-resistance determinant from one plasmid could become associated with another plasmid. Datta *et al.*[43] observed an apparently similar recombinational event that resulted in acquisition by the R64 plasmid of the TEM β-lactamase (Ap resistance) gene originally present on another plasmid, RP4; Richmond and Sykes[44] subsequently reported that the TEM β-lactamase gene derived from the related RP1 plasmid could integrate into the *E. coli* chromosome. Hedges and Jacob[45] isolated a series of plasmid derivatives that had received the Ap-resistance trait from RP4, and found that an increase in plasmid molecular weight accompanied the change to Ap resistance; other plasmids were then able to acquire the TEM β-lactamase resistance from the derivative plasmid. These authors concluded that the structural information for the TEM β-lactamase was carried by a transposable DNA sequence that they termed transposon A (TnA)[45]; Heffon *et al.*[46] subsequently demonstrated that the nucleotide sequences of the TnA segment, which is approximately 3×10^6 daltons in size, are common to naturally occurring plasmids of a variety of compatibility groups that specify the TEM type Ap resistance.

Although these recombinational events apparently involved movement of genetic information from one genome to another, they seemed to involve *rec*A-dependent general recombination. Other experimental evidence, however, has demonstrated that translocation of structurally defined segments of DNA carrying antibiotic resistant genes can occur by terminus-site-specific *rec*A-independent recombination. Some of these events involve known IS elements, which can also move from genome to genome independently of *rec*A function[47]; others seem to be associated with newly observed structurally defined segments of DNA that have been found on plasmids as inverted repeats.

During studies of mobilisation of a non-conjugative plasmid by a conjugally proficient Ap-resistance plasmid[19], a recombinant DNA molecule was isolated, which on further study was shown by electron microscope heteroduplex analysis to have been formed by translocation of a specific DNA segment from one plasmid to another[49]. Translocation of this DNA segment between plasmids did not require the bacterial *rec*A gene product, suggesting that it did not involve ordinary recombination mechanisms. The event was site specific so far as the transposable element itself is concerned (that is, the element could move between plasmids as a discrete structural unit), but the element could become inserted at different sites on the recipient pSC101 plasmid. Heteroduplex analysis showed that all of the sequences present in the translocating Ap-resistance segment were derived from a continuous region of the donor plasmid, and that insertion of the element into the recipient plasmid could occur in either direction[49]. Moreover, a 130-nucleotide long inverted repeat DNA sequence was present at the termini of the 4.5-kb transposable unit.

A schematic representation of these events is shown in Fig. 4. In effect, the plasmid DNA originally located at the ends of the inverted repeat termini of the transposable element is replaced by the recipient plasmid during the translocation event; it is not known, however, whether the transposable unit exists in a physically separate state during the process. In the hairpin-loop structure configuration, (Fig. 4a) the termini of the inverted repeats forming the stem of the Ap-resistance transposable segment are located around an axis of twofold rotational symmetry that marks the site of excision of the element from one plasmid and insertion of the element into another. Alternatively, such molecules may form loop configurations (Fig. 4b) similar to that proposed by Campbell[4] for integration of phage λ into the *E. coli* chromosome, or may form cruciform structures[50]. Although

the actual structural configuration for the transposable segment during the recombinational event is not known, it seems plausible that enzymes capable of recognising the inverted repeat termini of the element may be involved in translocation.

Subsequent experiments[51] have indicated that the structurally defined Ap-resistance element observed to translocate from the pSC50 plasmid on to plasmid pSC101[49] is indistinguishable by electron microscope heteroduplex analysis from the TnA element from R1-19 studied by others[46,52]. Heffron and coworkers[52] have shown that insertion of TnA into a plasmid results in polar mutations similar to those occurring as a consequence of insertion of IS units. Using heteroduplex mapping, these investigators[52] have also demonstrated that the TnA element can insert into multiple sites clustered in a region corresponding to only one-third of the length of the 5.5×10^6 dalton RSF1010 plasmid genome. Similar clustering of sites that have received TnA has been observed on the ColE1 plasmid[53] and on pSC101[51].

Other recent genetic and molecular studies have shown that TnA is one of a number of antibiotic resistance-gene segments of DNA that can undergo *rec*A-independent translocation. In 1972, Chan *et al.*[54] observed that growth of the *S. typhimurium* phage P22 in a bacterial strain carrying the R100 plasmid resulted in formation of an unusual phage varient that could transduce resistance to Tc at a high frequency. The Tc-resistant P22 phage seemed analogous to the antibiotic-resistant P22 phage observed by Dubnau and Stocker[34]; moreover, it did not seem to result from the mechanism normally responsible for the formation of P22 transducing phage, since insertion of Tc resistance into the phage genome did not occur at the site usually associated with attachment of the phage to the *Salmonella* chromosome.

The Tc-resistance DNA segment that had recombined with phage P22 was found to contain a hairpin-loop structure[55] similar to the one identified previously on the R6-5 plasmid[17]. Subsequent studies[56] have demonstrated that this structure, which has inverted repeats of the IS3 element at its termini[26], is capable of translocation from P22 on to many different sites of the *Salmonella* chromosome. Insertion of the transposable element into a structural gene abolishes the function of that gene; moreover, a polar effect on the expression of genes distal to the promoter occurs when the element is inserted into a group of genes that form a single transcription unit[56], as is characteristic of the IS elements[7]. The two inverted repeat copies of IS3 near the Tc-resistance gene seem to be able to undergo translocation concurrently and to carry along the intervening segment of DNA.

Insertion of the Tc-resistance hairpin-loop structure occurs without loss of genetic information from the molecule into which it inserts; most of the mutated bacteria formed as a consequence of insertion can revert to prototrophy[56]. However, excision of the transposible element is usually not precise. Although insertion of the Tc-resistance element into the *Salmonella* chromosome results in a distribution of auxotrophic mutants similar to that obtained by chemical mutagenesis[56], clustering of insertion sites within the histidine operon suggests that insertion may not be entirely random. As in the case of the TnA element, insertion and excision of the Tc-resistance (IS3) segment is independent of the function of the *rec*A gene of the host[56,57].

Berg *et al.*[58] have made analogous observations involving a segment of plasmid DNA that carries genes for resistance to the antibiotic kanamycin (Km). This segment contains an inverted repeat hairpin-loop structure similar in appearance to, but different in size from, the structure previously associated with Km resistance on various R plas-
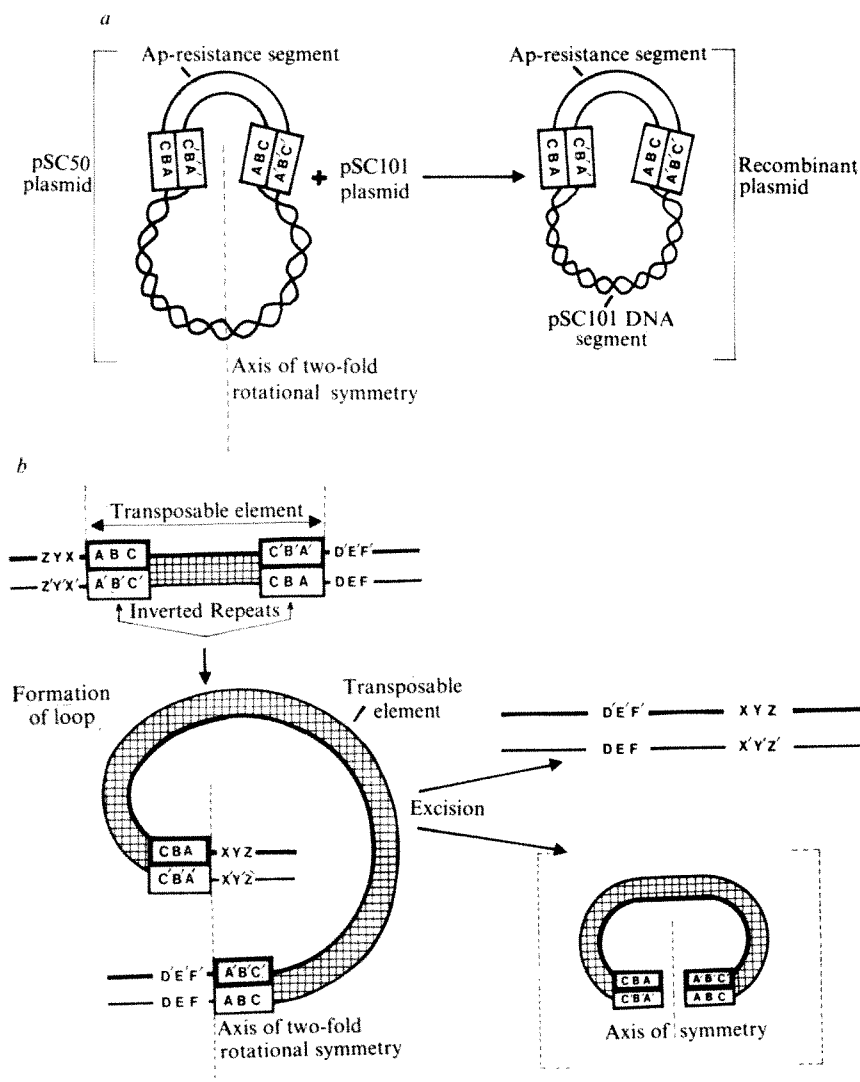
Fig. 4 Translocation of Ap-resistance DNA segment (TnA) from the pSC50 plasmid to pSC101. *a*, Schematic representation showing the Ap-resistance hairpin-loop structure with a 130–140 nucleotide-long inverted repeat stem. As a consequence of translocation, the pSC50 plasmid DNA at the base of the stem is replaced by the pSC101 plasmid (that is, the Ap-resistance segment is translocated from pSC50 to pSC101). Translocation of the hairpin-loop structure involves the ends of the palindromic (stem) sequences that bracket an axis of twofold rotational symmetry. The molecular nature of the DNA product resulting from excision of the Ap-resistance segment from pSC50 is not known. *b*, Possible Campbell model configuration for DNA segment involved in translocation. There is currently no evidence that the same event results in both translocation and restoration of DNA nucleotide sequence continuity at the site into which the Ap-resistance segment has been inserted; although excision with reversion to prototrophy, and translocation are both observed, these outcomes may represent separate events that occur on different DNA molecules, as appears to be the case with bacteriophage Mu[82,83]. The separated Ap-resistance transposable DNA segment is shown in brackets, since its existence as a separate unit has not been demonstrated.

mids[17]. The segment behaves functionally like the Tc-resistance and TnA elements; it can insert at several different sites on phage λ and excision leads to recovery of an intact phage genome. A second Km-resistance DNA segment lacking inverted repeat termini was also observed[38] to be capable of translocation from R plasmids to bacteriophage λ, suggesting that inverted repeats may not be essential for the translocation event or that one of the limbs of the inverted repeat is lost during translocation of this particular element.

Another translocating segment of DNA carrying an antibiotic resistance gene but lacking inverted repeats at its ends has been reported by Gottesman and Rosner[59]. This element had originally been acquired by *E. coli* bacteriophage P1 from the antibiotic resistance plasmid R10 (ref. 35). It was subsequently transferred to phage λ by a process shown to be independent of host *rec* or λ *red* recombination functions; however, possible involvement of P1 recombination functions in the translocation event was not excluded. Very recent investigations (MacHattie and Jackowski, personal communication) indicate that this Cm-resistance translocating segment of DNA contains directly repeated copies of the IS1 element at its termini.

During the studies of Kopecko and Cohen[49], a second transposable element that carries resistance to sulphonamide (Su), streptomycin (Sm) and Ap was observed to move from the pSC50 plasmid to pSC101. Subsequent experiments have shown that this 20.4-kb unit, which has been named TnS, contains the TnA element within it[51]. Inverted repeat segments, 130–140 nucleotides in length, are present at the termini of TnS as well as at the ends of TnA. More-

over, the termini of the TnS element seem to be highly interactive "hotspots" for additional illegitimate recombinational events involving the deletion or insertion of other segments of plasmid DNA[51]. Recently, a segment of DNA carrying genes for resistance to trimethoprim, streptomycin and spectinomycin has also been identified (TnC, ref. 60). Although it is not yet known whether this unit contains either a direct DNA repeat of inverted repeats at its terminal, it seems to be functionally similar to the other transposable antibiotic resistance units described above.

At this time, structurally defined DNA segments coding for resistance to ampicillin, streptomycin–spectinomycin, sulphonamide, chloramphenicol, tetracycline, and trimethoprim have been shown to be capable of translocation from genome to genome, and it seems that the sequential translocation of such segments may be the principal mechanism by which plasmids can accumulate multiple antibiotic resistance determinants. It seems likely that the widespread clinical and veterinary use of antimicrobial drugs has resulted in the biological selection of plasmids that acquire such transposable elements. Possibly, translocation of additional segments of DNA carrying various other plasmid or chromosomal genes remains undetected because comparable selective pressures have not been applied.

While inverted repeats have not been found on all of the antibiotic-resistance transposable DNA studied, they nevertheless seem to be involved prominently in the translocation process. In bracketing an axis of twofold rotational symmetry at the site of recombination of their termini, inverted repeats form a palindromic DNA sequence as noted above. Palindromes have also been observed at the

DNA interaction site for the *ter* enzyme of phage λ[61], at promoter and operator sites associated with λCI (ref. 62) and *E. coli lac*[63] operon repression, and at the sites of action of several restriction and modification enzymes[64,65], suggesting that they may provide highly specific recognition sites for protein–DNA interactions. Observations which show that certain palindrome-site-specific restriction endonucleases are coded by plasmid genes[64] raise the possibility that functionaly analogous plasmid-specified enzymes may have a role in recombinational events involving transposable elements. Moreover, recently obtained data[66] suggest that genetic information required for translocation of at least one of the antibiotic-resistance transposable elements (that is, TnA) is contained within the loop portion of the element itself. It remains to be determined whether other Tn units carry genes that specify products involved in translocation events.

Promotion of other illegitimate recombination events by transposable elements

Reif and Saedler[67] have reported that the IS1 element can promote *rec*A-independent deletion of DNA sequences immediately adjacent to the IS1 terminus. One endpoint of the deletion seems to be the terminus of IS1 itself, whereas the other endpoint is variable. Recent evidence from other laboratories suggests that transposable segments of DNA carrying antibiotic resistance genes can also promote deletion formation, and that such deletions occur at the termini of the inverted repeats. During heteroduplex analysis of the R100 plasmid, Hu *et al.*[29] observed that a transfer-deficient R100 derivative had deleted a DNA segment adjacent to the (IS3) terminus of the Tc-resistance hairpin-loop structure. Foster and Willets[68] subsequently found that following the loss of Tc resistance by R100-1, deletions extending into the adjacent *tra* (transfer) region of the plasmid are observed. Kleckner, Reichardt and Botstein (personal communication) have seen deletion formation also during excision of the Tc-resistance hairpin-loop structure from the *his* region of the *E. coli* chromosome, and D. Berg and B. Allet (personal communication) have observed small deletions in the λ phage chromosome following excision of a Km-resistance transposable element from this DNA segment.

The 130–140-nucleotide long inverted repeats at the termini of TnS also seem to be involved in deletion formation[51,69]. The pSC50 plasmid, which contains both the TnA and TnS translocating elements, is a Km-sensitive derivative of the Km-resistant R1-19 plasmid[49]. Heteroduplex analysis of pSC50 and its parent indicate that the entire segment of DNA located between a terminus of the TnS unit and a copy of IS1 is lost during the formation of pSC50, bringing the termini of TnS and IS1 in juxtaposition[51]. The mechanism by which transposable elements can promote deletion of adjacent DNA sequences is unknown. In at least some instances, however, deletion occurs while the transposable unit is *in situ*, and not at the time of excision or insertion of the element[69].

Conclusions

As a consequence of work carried out in a number of laboratories during the past several years, an understanding of the organisation and structural evolution of bacterial plasmids has begun to emerge. Terminus-site-specific recombination, which in *E. coli* and *Salmonella* is independent of bacterial *rec*A gene function, seems to have an important role; in fact, the addition of structurally-defined DNA segments on to plasmid, phage and bacterial chromosomes, and the excision of specific DNA segments from such chromosomes may represent the principal mechanism by which the organisation of prokaryotic (and perhaps eukaryotic) DNA has evolved. Certainly, this mechanism seems to account significantly for the structural and genetic diversity of plasmids, and to explain at least some types of interaction between extrachromosomal segments of DNA and the host cell chromosome.

Terminus-site-specific recombination also provides a potential mechanism for the exchange of genetic information between diverse biological species found in nature. There has been recent evidence that an Ap-resistance DNA segment found in *H. influenzae* is structurally and genetically similar to the TnA element isolated from *E. coli* plasmids[70]. In addition, translocation of prokaryotic gene segments of eukaryotic DNA sequences cloned in bacteria has been observed (A. C. Y. Chang and S. N. Cohen, unpublished data); it is tempting to speculate that prokaryotic-eukaryotic genetic recombination involving transposable genetic elements may occur under some circumstances in nature.

In addition to having an important role in the structural organisation of genes on prokaryotic genomes, transposable segments of DNA function as controlling elements or "switches" in turning on and off the expression of plasmid and chromosome genetic information. The remarkable dual role of these elements in affecting both the organisation and function of prokaryotic genomes is perhaps best illustrated by the Tc-resistance transposable segment. Inverted repeats of the IS3 element at the ends of this segment can accomplish movement of the intervening Tc-resistance gene between plasmids, chromosome, and phage, and can also lead to deletion of DNA sequences adjacent to the termini; another repeated copy of IS3 controls expression of the Tc-resistance gene contained within the segment of DNA that is translocated. The demonstration that still other insertion sequence elements carry promoter sites[25] for initiation of RNA synthesis provides another potential mechanism by which these units can function as moveable and invertable switches in the evolution of plasmid operons.

Parallels in the effects produced by transposable genetic elements in plasmids, and phenomena observed in eukaryotic systems suggest that similar DNA segments may have a role in the organisation and control of expression of genetic information in higher organisms. The transposable "controlling" elements that regulate phenotypic expression in maize[71-74] are known to be capable of movement to various sites on maize chomosomes; as a consequence of transposition to different locations, they can exert their controlling influence on a variety of genes. As is the case with the inverted repeats identified on transposable antibiotic resistance segments, transposable maize elements seem to promote deletions of various lengths which affect neighbouring genes. Insertion and excision of repeated DNA sequences has been shown recently also to be associated with genetic instability of the *white* locus on the X chromosome of *Drosophila melanogaster*[75,76]. The mutations produced in this locus can be transposed to another linkage group, leaving behind a genetic deletion. Moreover, foldback DNA characteristic of inverted repeats has been found in various eukaryotic chromosomes, and visualisation of this DNA in the electron microscope shows hairpin-loop structures similar to those seen on plasmids[77-79].

Analogies between phenomena observed in eukaryotic systems and the transposable genetic elements and inverted repeats identified on plasmids are presently only speculative; nevertheless, bacterial plasmids may prove to be a highly useful model system for elucidating mechanisms involving the organisation and control of genetic information on the complex chromosomes of higher organisms.

1 Cozzarelli, N. R., Kelly, R. B., and Kornberg, A., *Proc. natn. Acad. Sci. U.S.A.*, 69, 2110–2114 (1968).
2 Low, K. B., *Bact. Rev.*, 36, 587–607 (1972).
3 Clark, A. J., *A. Rev. Genet.*, 7, 67–86 (1973).
4 Campbell, A., *Adv. Genet.*, 11, 101–145 (1962).
5 Franklin, N., in *The Bacteriophage Lambda* (edit. by Hershey, A. D.), 175 (Cold Spring Harbor Laboratory, New York, 1971).
6 Cohen, S. N., and Kopecko, D. J., *Fedn Proc.*, 35, 2031–2036 (1976).
7 Starlinger, P., and Saedler, H., *Crit. Rev. Microbiol.* (in the press).
8 Nisioka, T., Mitani, M., and Clowes, R., *J. Bact.*, 97, 376–385 (1969).
9 Cohen, S. N., and Miller, C. A., *J. molec. Biol.*, 50, 671–687 (1970).
10 Haapala, D. K., and Falkow, S., *J. Bact.*, 106, 294–215 (1971).
11 Kopecko, D. J., and Punch, J. D., *Ann. N.Y. Acad. Sci.*, 182, 207–216 (1971).
12 Rownd, R., and Mickel, S., *Nature new Biol.*, 234, 40–43 (1971).
13 Clowes, R. C., *Bact. Rev.*, 36, 361–405 (1972).
14 Rownd, R. H., and Perlman, D., *Nature*, 259, 281–284 (1976).
15 Cohen, S. N., and Miller, C. A., *Proc. natn. Acad. Sci. U.S.A.*, 67, 510–516 (1970).
16 Silver, R. P., and Falkow, S., *J. Bact.*, 104, 331–339 (1970).
17 Sharp, P. A., Cohen, S. N., and Davidson, N., *J. molec. Biol.*, 75, 235–255 (1973).
18 Sharp, P. A., Hsu, M. T., Ohtsubo, E., and Davidson, N., *Proc. natn. Acad. Sci. U.S.A.*, 71, 471–497 (1973).
19 Cohen, S. N., and Chang, A. C. Y., *Proc. natn. Acad. Sci. U.S.A.*, 70, 1293–1297 (1973).
20 Cohen, S. N., Chang, A. C. Y., and Hsu, L., *Proc. natn. Acad. Sci. U.S.A.*, 69, 2110–2114 (1972).
21 Malamy, M. H., *Cold Spring Harb. Symp. quant. Biol.*, 31, 189–201 (1966).
22 Saedler, H., and Starlinger, P., *Molec. gen. Genet.*, 100, 178–189 (1967).
23 Jordan, E., Saedler, H., and Starlinger, P., *Molec. gen. Genet.*, 100, 299–306 (1968).
24 Shapiro, J. A., *J. molec. Biol.*, 40, 93–105 (1969).
25 Saedler, H., Reif, H. J., Hu, S., and Davidson, N., *Molec. gen. Genet.*, 132, 265–289 (1974).
26 Ptashne, K., and Cohen, S. N., *J. Bact.*, 122, 776–781 (1975).
27 Fiandt, M., Szybalski, W., and Malamy, M. H., *Molec. gen. Genet.*, 119, 223–231 (1972).
28 Michaelis, G., Saedler, H., Venkov, P., and Starlinger, P., *Molec. gen. Genet.*, 104, 371–377 (1969).
29 Hu, S., Ohtsubo, E., Davidson, N., and Saedler, H., *J. Bact.*, 122, 764–775 (1975).
30 Rownd, R. H., Perlman, D., and Goto, N., in *Microbiology 1974* (American Society for Microbiology, Washington DC, 1975).
31 Hu, S., Ohtsubo, E., and Davidson, N., *J. Bact.*, 122, 749–763 (1975).
32 Davidson, N., Deonier, R. C., Hu, S., and Ohtsubo, E., in *Microbiology, 1974* (edit. by Schlesenger, D.), 56 (American Society for Microbiology, Washington DC, 1975).
33 Hu, S., Ptashne, K., Cohen, S. N., and Davidson, N., *J. Bact.*, 123, 687–692 (1975).
34 Dubnau, E., and Stocker, B. A. D., *Nature*, 204, 1112–1113 (1974).
35 Kondo, E., and Mitsuhashi, S., *J. Bact.*, 88, 1266–1276 (1964).
36 Starlinger, P. et al., *Molec. gen. Genet.*, 122, 279–286 (1973).
37 Iyobe, S., Hashimoto, H., and Mitsuhashi, S., *Jap. J. Microbiol.*, 13, 225–232 (1969).
38 Iyobe, S., Hashimoto, H., and Mitsuhashi, S., *Jap. J. Microbiol.*, 14, 463–471 (1970).
39 Harada, K., Kameda, M., Suzuki, M., Shigebara, S., and Mitsuhashi, S., *J. Bact.*, 93, 1236–1241 (1967).
40 Kameda, M., Harada, K., Suzuki, M., and Mitsuhashi, S., *Jap. J. Microbiol.*, 13, 255–262 (1969).
41 Mitsuhashi, S., *J. Infect. Dis.*, 119, 89–100 (1969).
42 Anderson, E. S., Keleman, M. V., Jones, C. M., and Pitton, J. S., *Genet. Res. Camb.*, 11, 119–124 (1968).

43 Datta, N., Hedges, R. W., Shaw, E. J., Sykes, R. B., and Richmond, M. H., *J. Bact.*, 108, 1244–1249 (1971).
44 Richmond, M. H., and Sykes, R. B., *Genet. Res.*, 20, 231–237 (1972).
45 Hedges, R. W., and Jacob, A. E., *Molec. gen. Genet.*, 132, 31–40 (1974).
46 Heffron, F., Sublett, R., Hedges, R. W., Jacob, A., and Falkow, S., *J. Bact.*, 122, 250–256 (1975).
47 Jaskunas, S. R., Lindahl, L., Nomura, M., and Burgess, R. R., *Nature*, 257, 458–462 (1975).
48 Bennett, P. M., and Richmond, M. H., *J. Bact.*, 126, 1–6 (1976).
49 Kopecko, D. J., and Cohen, S. N., *Proc. natn. Acad. Sci. U.S.A.*, 72, 1373–1377 (1975).
50 Sobell, H. M., *Proc. natn. Acad. Sci. U.S.A.*, 69, 2483–2487 (1972).
51 Kopecko, D. J., Brevet, J., and Cohen, S. N. (in the press).
52 Heffron, F., Rubens, C., and Falkow, S., *Proc. natn. Acad. Sci. U.S.A.*, 72, 3623–3627 (1975).
53 So, M., Gill, R., and Falkow, S., *Molec. gen. Genet.*, 142, 239–249 (1975).
54 Chan, R. K., Botstein, T., Watanabe, T., and Ogata, Y., *Virology*, 50, 883–898 (1972).
55 Tye, B. K., Chan, R. K., and Botstein, D., *J. molec. Biol.*, 85, 485–500 (1974).
56 Kleckner, N., Chan, R., Tye, B., and Botstein, D., *J. molec. Biol.*, 97, 561–575 (1975).
57 Foster, T. J., Howe, T. G. B., and Richmond, K. M. V., *J. Bact.*, 124, 1153–1158, (1975).
58 Berg, D. E., Davies, J., Allet, B., and Rochaix, J., *Proc. natn. Acad. Sci. U.S.A.*, 72, 3628–3632 (1975).
59 Gottesman, M., and Rosner, J. L., *Proc. natn. Acad. Sci. U.S.A.*, 72, 5041–5045 (1975).
60 Barth, P. T., Datta, N., Hedges, R. W., and Grintner, N. J., *J. Bact.*, 125, 800–810 (1976).
61 Weigel, P. H., England, P. T., Murry, K., and Old, R. W., *Proc. natn. Acad. Sci. U.S.A.*, 70, 1151–1155 (1973).
62 Maniatis, T., Ptashne, M., Barrell, B. G., and Donelson, J., *Nature*, 250, 394–397 (1974).
63 Gilbert, W., and Maxam, A., *Proc. natn. Acad. Sci. U.S.A.*, 70, 3581–3584 (1973).
64 Boyer, H. W., *Fedn Proc.*, 33, 1125–1127 (1974).
65 Kelly, T. J., and Smith, H. O., *J. molec. Biol.*, 51, 393–409 (1970).
66 Heffron, F., Bedinger, P., Champoux, J., and Falkow, S. (in the press).
67 Reif, J. H., and Saedler, H., *Molec. gen. Genet.*, 137, 17–28 (1975).
68 Foster, T. J., and Willets, N., *J. gen. Microbiol.*, 93, 133–140 (1976).
69 Nisen, P., Koepcko, D. J., and Cohen, S. N. (in the press).
70 de Graff, J., Elwell, L. P. and Falkow, S., *J. Bact.*, 126, 439–446 (1976).
71 McClintock, B., *Cold Spring Harb. Symp. quant. Biol.*, 21, 197–216 (1956).
72 McClintock, B., *Brookhaven Symposia in Biology*, 18, 162–184 (1965).
73 Peterson, P. A., *Genetica*, 41, 33–56 (1970).
74 Fincham, J. R. S., and Sastry, G. R. K., *A. Rev. Genet.*, 8, 15–50 (1974).
75 Green, M. M., *Genetics*, 61, 423–428 (1969).
76 Rasmuson, B., Green, M. M., and Karlsson, B., *Molec. gen. Genet.*, 133, 237–247 (1974).
77 Wilson, D. A., and Thomas, C. A., *molec. Biol.*, 84, 115–144 (1974).
78 Schmid, C., Manning, J., and Davidson, N., *Cell*, 5, 159–172 (1975).
79 Cech, T., and Hearst, J., *Cell*, 5, 429–446 (1975).
80 Cohen, S. N., Chang, A. C. Y., Boyer, H. W., and Helling, R. B., *Proc. natn. Acad. Sci. U.S.A.*, 70, 3240–3244 (1973).
81 Davis, R. W., Simon, M. N., and Davidson, N., in *Methods in Enzymology*, 21, 413–428 (Academic, New York, 1971).
82 Bukhari, A., *J. molec. Biol.*, 96, 87–99 (1975).
83 Howe, M., and Bade, E., *Science*, 190, 624–632 (1975).

articles

Fission-track dating of pumice from the KBS Tuff, East Rudolf, Kenya

Anthony J. Hurford & A. J. W. Gleadow*

Department of Geology, Birkbeck College, University of London, 7/15 Gresse Street, London W1P 1PA, UK

C. W. Naeser

US Geological Survey, Federal Center, Denver, Colorado 80225

Fission-track dating of zircon separated from two pumice samples from the KBS Tuff in the Koobi Fora Formation, in Area 131, East Rudolf, Kenya, gives an age of 2.44 ± 0.08 Myr for the eruption of the pumice. This result is compatible with the previously published K–Ar and ⁴⁰Ar/³⁹Ar age spectrum estimate of 2.61 ± 0.26 Myr for the KBS Tuff in Area 105, but differs from the more recently published K–Ar date of 1.82 ± 0.04 Myr for the KBS Tuff in Area 131. This study does not support the suggestion that pumice cobbles of different ages occur in the KBS Tuff.

THE Plio-Pleistocene sedimentary sequence in the East Rudolf basin of Northern Kenya, has been the site of many important hominid[1], archaeological[2] and palaeontological finds[3]. Intercalated with the sequence of lacustrine, deltaic and fluviatile sediments are a series of prominent felsic volcanic tuffs. These volcanic units have been used to establish a time framework for the rocks of the East Rudolf Basin[4]. K–Ar and ⁴⁰Ar/³⁹Ar dating techniques have been applied to >100 rock and mineral samples from East Rudolf, but interpretation of the dates determined by these methods has not been straightforward. Geological[5] and

*Present address: Department of Geology, School of Earth Sciences, University of Melbourne, Parkville, Victoria, Australia.

analytical[8,9] factors have been postulated to explain the scatter of K–Ar and [40]Ar/[39]Ar apparent ages obtained from volcanic sanidine–anorthoclase crystals separated from pumice cobbles in the tuffs.

The KBS Tuff, lying within the Lower Member of the Koobi Fora formation, has attracted considerable interest amongst palaeontologists and anthropologists because it contains an important stone artefact site[2] and overlies the bed which contained the hominid cranium KNM ER 1470 (ref. 1). In 1970, Fitch and Miller[6] published a K–Ar and [40]Ar/[39]Ar age spectrum analysis of fresh sanidine crystals separated from pumice cobbles present in the KBS Tuff in Area 105, giving a preliminary date of 2.61 ± 0.26 Myr for the KBS Tuff horizon. Recomputation of this data using improved [40]Ar/[39]Ar techniques, together with additional [40]Ar/[39]Ar dating studies leads Fitch *et al.*[7] to the conclusion that a more precise minimum age for the KBS Tuff is 2.42 ± 0.01 Myr.

Curtis has described the original 2.61 ± 0.26 Myr date for the KBS Tuff as being much questioned in private anthropological and palaeontological circles[8]. These doubts led Curtis *et al.* to undertake an independent check, using conventional K–Ar dating techniques on sanidine and glass separated from KBS pumice[9]. The tuff known as KBS in Areas 10 and 105 is distinguished by Curtis *et al.*[9] from the tuff known as KBS in Area 131 on the basis of K–Ar total fusion dates obtained from these areas: 10 and 105 give a mean date of 1.60 ± 0.05 Myr, 131 giving a mean date of 1.82 ± 0.04 Myr. The 2.61 ± 0.26-Myr date of Fitch and Miller[6] is explained by Curtis *et al.*[9] as possibly resulting from the presence of older pumice blocks in the KBS Tuff. Curtis underlines the desirability of checking radiometric dates by the use of two different methods, and he cites the fission-track dating work on the East Rudolf Tuffs being undertaken at Birkbeck College[8]. Because of the great interest in the KBS Tuff horizon and the apparent conflict between the K–Ar and [40]Ar/[39]Ar results, we present here the first dates obtained using the independent technique of fission-track dating on zircon from pumice cobbles collected at one locality in the KBS Tuff. This work represents part of a larger fission-track dating study being made on each of the tuff horizons in the East Rudolf Basin. Full publication of the fission-track results will be made on completion of the project. The theory and techniques of fission-track dating have been comprehensively described[10], as has the specific use of zircon in dating young volcanic ash beds[11].

Method

The two pumice samples used in this study, FMA 517 and ER 74/131 came from within 5 m of each other, from the KBS Tuff in Area 131. Sample FMA 517 was collected in 1974 by I. C. Findlater, F. J. Fitch and A. J. Hurford, and was used in a further [40]Ar/[39]Ar dating study undertaken at Cambridge[7]. FMA 517 consists of several well rounded

pumice lumps collected within 4 m of each other. The pumice is rich in euhedral feldspar phenocrysts and contains glass showing partial replacement by calcite. Sample ER 74/131 was collected by T. E. Cerling and was used in the K–Ar dating study undertaken at Berkeley[9]. This single pumice lump contained dark, isotropic glass and abundant euhedral feldspar crystals. The pumice samples were crushed and carefully digested in a mixture of concentrated HF and concentrated H_2SO_4 for 7 d (ref. 16). This digestion process dissolved most of the rock leaving a concentrate of zircon which was further purified by conventional mineral separation techniques. The zircons consist of two distinct varieties: ~95% are colourless, water-clear unzoned euhedral crystals, 200–50 μm in length and often containing many randomly orientated acicular inclusions together with irregularly shaped fluid inclusions. These crystals all have very low spontaneous fission-track densities. From the perfect, euhedral nature of these zircons, they must be regarded as juvenile crystals, primary to the eruption of the enclosing pumice. Approximately 5% of the zircons are variable in form and much darker in colour. They are light brown to red, often strongly zoned and have a very rounded form, with no discernible crystal faces or inclusions. Since these crystals all have high spontaneous fission-track densities, they appear to be old, detrital zircons and therefore they are excluded from the determinations shown in Table 1.

Four determinations were carried out on sample FMA 517, three at Birkbeck College, London, and one at the US Geological Survey, Denver, Colorado. One determination was carried out on sample ER 74/131 at Birkbeck College. The techniques used for dating the zircons in both laboratories are essentially those described by Naeser[12], except that the tracks were etched in a molten eutectic of KOH and NaOH at 220 °C for between 60 and 100 h (ref. 13). The external detector method was used, [238]U spontaneous fission tracks being counted in the zircon and [235]U induced fission tracks counted in a calibrated external mica detector. Neutron irradiations were carefully monitored by counting tracks from the NBS standard glass SRM 612, which had been calibrated previously against the NBS dosimeter glass SRM 962 and a large number of cobalt wire activation measurements. Using these techniques and a value for the [238]U spontaneous fission decay constant, λ_f of 6.85×10^{-17} yr^{-1} (ref. 10) we have obtained ages on standard zircons which agree very closely with their independently known ages[13]. We believe, therefore, that this value of λ_f is that which is most appropriate to fission-track dating. Fission-track ages have been calculated using the equation of Fleischer *et al.*[14]. The results of the five determinations are shown in Table 1.

In addition to the dates shown in Table 1, counting of three of the detrital grains from FMA 517 by one of us (AG) gave dates of 303 ± 32, 380 ± 62 and 293 ± 52 Myr,

Table 1 Fission-track dating results of two pumice samples from the KBS Tuff, East Rudolf

Sample no. and Determination no.	No. of zircons	Spontaneous tracks No. counted	ρ_s (cm^{-2})	Induced tracks No. counted	ρ_i (cm^{-2})	Neutron fluence φ (n cm^{-2})	Age (±1σ)(Myr)
FMA 517 (1) AJH	10	47	7.05×10^4	1,428	4.28×10^6	2.52×10^{15}	2.55 ± 0.38
FMA 517 (2) AJH	10	56	7.87×10^4	1,129	3.18×10^6	1.59×10^{15}	2.42 ± 0.32
FMA 517 (3) AJWG	6	36	7.09×10^4	802	3.32×10^6	1.81×10^{15}	2.37 ± 0.41
FMA 517 (4) CWN	12	22	9.55×10^4	311	2.70×10^6	1.11×10^{15}	2.41 ± 0.53
						Mean date of FMA 517	2.44 ± 0.08
ER 74/131 (1) AJH	8	21	1.34×10^5	411	5.26×10^6	1.55×10^{15}	2.43 ± 0.55

suggesting an age for the detrital component of about 325 Myr. These results confirm that the pumice blocks had been contaminated by older zircons, possibly during eruption, or by the penetration of detrital grains into the pumice vesicles, during transportation or on deposition. During the fission-track dating procedure this inherited component can be identified easily by the morphology and high track density of these particular zircons. For this reason, it is quite certain that none of these older crystals has been included in the dating of the primary zircons.

The mean of the four determinations on FMA 517 is 2.44 ± 0.08 Myr and this cannot be distinguished from the single determination of 2.43 ± 0.55 Myr for ER 74/131. These dating results provide strong evidence that the primary zircon component in pumice samples from the KBS Tuff in Area 131 has an age close to 2.44 Myr. This may be taken also as the age of eruption and deposition of the KBS Tuff, provided the zircon dates have not been reset by some later event and provided the dated pumices are of the same age as the surrounding tuffaceous matrix. Extrapolation of laboratory annealing studies[14,16] indicates that fission tracks in zircon should be stable over a period of 3 Myr at temperatures up to at least 300 °C. It is highly unlikely, therefore, that the zircon fission-track dates have been reset by any later thermal events.

Curtis *et al.*[9] report that sample ER 74/131 has a K–Ar apparent age of 1.82 ± 0.04 Myr whereas Fitch *et al.*[7] interpret a ⁴⁰Ar/³⁹Ar age spectrum on FMA 517 as indicating an original date of 2.47 ± 0.60 Myr, overprinted at 1.91 ± 0.03 Myr. To explain such variation in K–Ar and ⁴⁰Ar/³⁹Ar estimates of the age of the KBS Tuff, it has been suggested that pumice of more than one age may be present in this horizon[9]. Our limited data do not support this conclusion. The pumice samples ER 74/131 and FMA 517 dated in this study give virtually identical fission-track results of ~2.44 Myr, suggesting that these two samples of pumice cobbles are of the same age.

Conclusion

The mean zircon fission-track date of 2.44 ± 0.08 Myr provides a close estimate of the age of the eruption of the pumice samples collected from the KBS Tuff in Area 131. These fission-track results are also consistent with the palaeomagnetic stratigraphy established by Brock and Isaac[17] on the basis of the original 2.61-Myr date[6] for the KBS Tuff. From consideration of the data currently available, the date of 2.44 ± 0.08 Myr probably represents a close estimate of the age of eruption and deposition of the KBS Tuff.

1 Leakey, R. E. F., *Nature*, **242**, 170–173 (1973).
2 Isaac, G. Ll., Leakey, R. E. F., and Behrensmeyer, A. K., *Science*, **173**, 1129 (1971).
3 Maglio, V. J., *Nature*, **239**, 379–385 (1972).
4 Fitch, F. J., Findlater, I. C., Watkins, R. T., and Miller, J. A., *Nature*, **251**, 213–215 (1974).
5 Fitch, F. J., and Miller, J. A., in *Stratigraphy, Palaeoecology and Evolution in the Lake Rudolf Basin* (edit. by Coppens, Y., Isaac, G. Ll., and Howell, F. C.), (University of Chicago Press, Chicago, 1975).
6 Fitch, F. J., and Miller, J. A., *Nature*, **226**, 226–228 (1970).
7 Fitch, F. J., Hooker, P. J., and Miller, J. A., *Nature*, **263**, 740–744 (1976).
8 Curtis, G. H., *World Archaeol.*, **7**, 198–208 (1975).
9 Curtis, G. H., Drake, R. F., Cerling, T. E., Cerling, B. L., and Hampel, J. H., *Nature*, **258**, 395–398 (1975).
10 Fleischer, R. L., Price, P. B., and Walker, R. M., *Nuclear Tracks in Solids* (University of California Press, 1975).
11 Naeser, C. W., Izett, G. A., and Wilcox, R. E., *Geology*, **1**, 187–189 (1973).
12 Naeser, C. W., *Science*, **165**, 388 (1969).
13 Gleadow, A. J. W., Hurford, A. J., and Quaife, R., *Earth planet. Sci. Lett.* (in the press).
14 Fleischer, R. L., Price, P. B., and Walker, R. M., *J. geophys. Res.*, **70**, 1497–1502 (1965).
15 Fleischer, R. L., and Price, P. B., *Phys. Rev.*, **133**, B63–64 (1964).
16 Krishnaswami, S., Lal, D., and Prabhu, N., *Earth planet. Sci. Lett.*, **22**, 51–59 (1974).
17 Brock, A., and Isaac, G. Ll., *Nature*, **247**, 344–348 (1974).

⁴⁰Ar/³⁹Ar dating of the KBS Tuff in Koobi Fora Formation, East Rudolf, Kenya

F. J. Fitch
Department of Geology, Birkbeck College, University of London, 7/15 Gresse Street, London W1P 1PA, UK

P. J. Hooker & J. A. Miller
Department of Geodesy and Geophysics, University of Cambridge, Madingley Rise, Madingley Road, Cambridge CB3 0EZ, UK

⁴⁰Ar/³⁹Ar and conventional K–Ar age determinations on volcanic sanidine–anorthoclase crystal concentrates from pumice lumps contained within the KBS Tuff in the Koobi Fora Formation at East Rudolf indicate an apparent age near 2.42 Myr. Recomputation of data obtained in 1969 from this tuff suggests a more accurate apparent isochron age of 2.42 ± 0.01 Myr for the feldspar, coincident with total fusion conventional K–Ar and ⁴⁰Ar/³⁹Ar dates of ~2.4 Myr from the same sample. Age spectra and total fusion dates obtained from other samples of the KBS Tuff confirm this conclusion. Thus, apparent ages of < 2.4 Myr derived from some of our samples and conventional total fusion K–Ar apparent ages in the range 1.6–1.8 Myr obtained from the KBS Tuff by other workers are regarded as discrepant, and may have been obtained from samples affected by argon loss. This latter conclusion is compatible with our ₚsector interpretation of the age spectrum analyses of KBS pumice feldspars. Argument for there being more than one tuff complex or an admixture of older pumice in the KBS Tuff is not supported by our work.

CONTROVERSY over the age of the KBS Tuff in the Koobi Fora Formation of the East Rudolf sedimentary basin of northern Kenya has arisen very largely from apparently conflicting evidence derived from the different approaches to the measurement of geological time. Exploration of the basin in 1968 and 1969 led to the identification of the KBS Tuff and an overlying unconformity[1–3]. Samples were collected then from an outcrop of this tuff in Area 105, East Rudolf and forwarded to Cambridge for dating[4]. It is now known that the KBS Tuff was deposited by a major river entering Lake Turkana (formerly Lake Rudolf) from the north-east, and numerous lines of evidence suggest that the time between eruption and deposition was negligible[5–7]. Exploratory conventional K–Ar dating in 1969 revealed the presence of detrital impurities in the vitric tuff sample that made its apparent ages discrepantly high (> 200 Myr). It was immediately clear that unless contamination-free fractions could be found the East Rudolf tuffs would not produce satisfactory samples for geochronometric analysis. Attention was concentrated, therefore, on pumice samples and on a large clean sample of volcanic feldspar extracted from pumice in the field. These

samples were carefully prepared in the laboratory to avoid contamination and analysed by the conventional K–Ar total fusion method and by the then new ^{40}Ar/^{39}Ar total fusion and age spectrum methods[8-11]. Developments in the analytical techniques of ^{40}Ar/^{39}Ar dating since then[12-14] enable recomputation of the results obtained, using, in
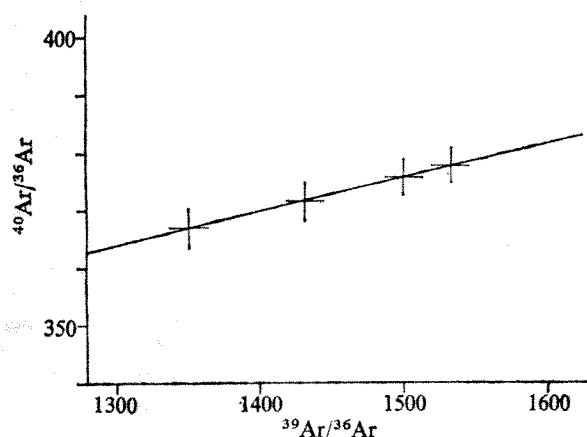
Fig. 1 Isochron diagram for the plateau sector of age spectrum No. 1 (Fig. 1 of ref. 4) of subsample Leakey I/B2 (d). Apparent age of isochron is 2.42 ± 0.01 Myr; intercept value on ^{40}Ar/^{36}Ar axis is 294.

addition, a more accurate value for the constant of proportionality (J) used in the 1969 experiments. These revised data are listed in Table 1 along with the total fusion K–Ar ages obtained from other subsamples of sample Leakey I for comparison.

Detrital contamination

In all K–Ar dating, some argon loss must be suspected in total fusion results until disproved. In addition, small contamination errors may be present in our total rock pumice samples even though they were selected from the cores of well-washed larger lumps. The concentrates of volcanic feldspar sent to the UK in 1969 consisted of a small number of large (up to 7 mm average diameter) euhedral sanidine–anorthoclase crystals and a greater quantity of small (1.0–5.0 mm average diameter) euhedral feldspar crystals and aggregates containing aegerine–augite and/or iron ore as well as feldspar. A number of different (unequal) subsamples were selected from this feldspar concentrate for dating, so that the presence of any 'detrital' contaminent would be revealed by variation in their apparent ages. The largest, most euhedral feldspars, were hand-picked for ^{40}Ar/^{39}Ar age spectrum analysis No. 1.

Five further subsamples, each consisting of smaller feldspars free from pyroxene and iron ore, were prepared and three of these were subsequently used for total fusion ^{40}Ar/^{39}Ar and total fusion conventional K–Ar dating.

The virtual coincidence of the apparent ages obtained from determinations made by different methods on four different subsamples of the feldspar concentrate Leakey I/B2 suggests that this concentrate does not suffer from significant detrital contamination. Error attributable to variable but very minor argon loss remains, however, and this is confirmed by the shape of the age spectrum. Thus, the best minimum age for the KBS Tuff currently available from this sample is 2.42 ± 0.01 Myr (revised apparent isochron age of plateau sector of age spectrum No. 1).

Argon loss

During the period 1971–73 further samples were collected from the KBS Tuff at various localities in the East Rudolf basin by I. C. Findlater. Each of the samples was a cleaned and acid-washed feldspar crystal concentrate extracted from
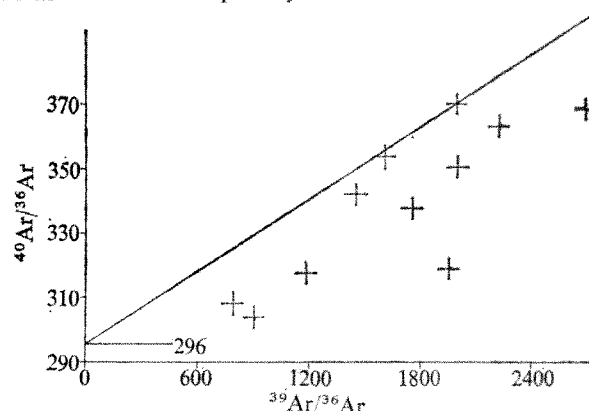
Fig. 2 Correlation diagram plot of total fusion ^{40}Ar/^{39}Ar age determination data obtained from feldspar concentrates separated from KBS pumice, East Rudolf, Kenya[15]. ^{39}Ar/^{36}Ar values are normalised to a constant J. See text. The straight line is the 2.4-Myr isochron envelope.

pumice by hand picking only the most obviously juvenile crystals from sawn slabs of pumice beneath a binocular microscope. The results of twelve total fusion and six age spectrum ^{40}Ar/^{39}Ar age determinations on these KBS feldspar samples were briefly summarised at a symposium held in Nairobi in 1973[15]. The apparent ages obtained from samples from Areas 10, 130, 131 and 105 show a scatter between 0.5 and 2.4 Myr. The available ^{40}Ar/^{39}Ar total fusion data from the KBS Tuff pumice feldspars is plotted as a correlation diagram in Fig. 2.

Because of the care taken in their preparation, and by analogy with the data obtained from the various Leakey

Table 1 Conventional K–Ar, new and revised ^{40}Ar/^{39}Ar dating results from sample Leakey I, KBS Tuff, Area 105, East Rudolf, Kenya (Q = 0.15152 GG2 muscovite monitor)

	Total fusion K–Ar dates (Myr)	Total fusion ^{40}Ar/^{39}Ar dates (Myr)	^{40}Ar/^{39}Ar age spectrum dates
Leakey I/B1 four pumice cores, total rock	3.63 ± 2.1* 2.40 ± 1.0	3.12 ± 1.1*† 2.26 ± 0.5†	
Leakey I/B2 feldspar crystal concentrate (six independent subsamples: a–f).	(a) 2.38 ± 0.3 (b) 2.37 ± 0.3	(c) 2.39 ± 0.3†	(d) 2.42 ± 0.01 Myr. Four point plateau sector analysis of age spectrum No. 1 (see Fig. 1)

*Contamination may have been present in these two pumice total rock samples, but in view of the large experimental errors inherent in their analysis, this cannot be regarded as proven.
†Revision of previously published data.
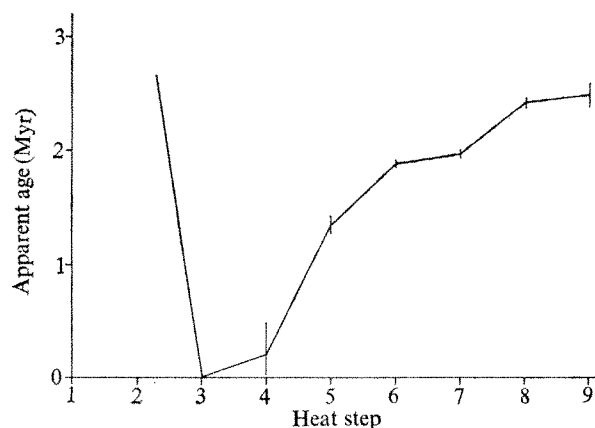
Fig. 3 $^{40}Ar/^{39}Ar$ age spectrum of KBS pumice feldspar sample FMA 274, Area 10, East Rudolf. The two plateau sectors of the spectrum have apparent ages of 1.94 Myr and 2.46 Myr respectively.

1/B2 feldspar crystal subsamples, the results are very unlikely to be marred by contamination. The scatter of apparent ages obtained must be explained, therefore, by the loss of varying amounts of Ar from the feldspars. If it is accepted that contamination is not significant in these samples, then a minimum age of 2.4 Myr for the KBS Tuff is indicated assuming an intercept value of 296 on the $^{40}Ar/^{36}Ar$ axis.

High temperature volcanic sanidine–anorthoclases are highly unstable minerals with aluminium and alkali atoms at random sites in the lattice. During subsequent changes of state towards more ordered low temperature forms of greater stability, significant losses of Ar may result. If these order–disorder processes are accentuated at any time by periods of thermal or hydrothermal activity, then related Ar loss may be identifiable in the K–Ar age patterns of samples of these feldspars. We believe that the complexities of some of the age spectra we have obtained from KBS feldspar samples can thus be explained and may even be resolvable[5,15].

Age spectra

Complex age spectra from four samples of KBS pumice feldspar from Area 130, East Rudolf (FMA 201, 206, 225, 227) provide strong evidence of K–Ar age discrepancy[16,17], which we attribute to argon loss. Two major age components can be recognised in the four age spectra: one apparently at 1.75–1.8 Myr and another apparently at ~ 1.07 Myr. When the data defining these sectors are plotted on correlation diagrams, regression line intercept values considerably less than 296 are obtained on the $^{40}Ar/^{36}Ar$ axis, thus giving further support to the overprinting hypothesis[17,18]. Similarly it can be suggested that at least two age sectors, ~ 1.9 Myr and 1.02 Myr are present in the complex age spectrum obtained from a further sample (FMA 294) of KBS pumice feldspar from Area 105. Such suggestions are not, however, conclusive.

An age spectrum obtained from a sample of clean, detritus-free KBS pumice feldspar (FMA 274) from Area 10, East Rudolf is, however, particularly instructive. Two aspects of the data from this sample are illustrated in Figs 3 and 4. The age spectrum (Fig. 3) is seen to have two short plateau segments apparently ~1.94 Myr and 2.46 Myr respectively. This type of double plateau age spectrum might be interpreted as suggesting either partial overprinting at 1.94 Myr of feldspar originally 2.46 Myr old, or the presence of contamination, but when the data are examined on a correlation diagram (Fig. 4) it is seen that another interpretation is possible: two slightly different feldspar components of unlike initial argon composition

but of identical apparent age are present, one having a regression age of 2.39±0.01 Myr (intercept value, 297) the other 2.42±0.03 Myr (intercept value, 272). The difference in intercept value between these two age components could result from one being accessory feldspar outgassed during the eruptive event of 2.4 Myr ago, whereas the other is of entirely juvenile origin. We feel that the data from sample FMA 274 provide strong additional evidence for a 2.4-Myr age for the KBS Tuff. Thus, our dating of samples of KBS pumice feldspar from other localities at East Rudolf seems to confirm the minimum age of 2.42 Myr derived from sample Leakey I and to suggest that discrepancies from Ar loss, particularly ~ 1.75–1.9 Myr and 1.05 Myr may be important in many KBS samples (see also the data from sample FMA 517, discussed below).

Opposition to our dating

Over the past five years, opposition to the acceptance of a ~2.5-Myr age for the KBS Tuff has come from three sources: first, archaeologists and palaeoanthropologists disturbed by the consequent antiquity of hominid fossils and stone tools found close to or associated with the KBS Tuff, second, palaeontologists reporting apparent misfits between the faunal sequences at East Rudolf and elsewhere, and third, from a small programme of conventional total fusion K–Ar age determinations on East Rudolf pumice samples undertaken at Berkeley[19-21]. We believe that the attempts at biostratigraphic and archaeological correlation so far made have been premature and confidently expect the problems raised by these workers[22-24] to disappear as more accurate faunal lists and local time scales are produced. The inconsistencies that have arisen from the application of different K–Ar dating techniques to samples of the same tuff do, however, require further comment now.

At present, the KBS Tuff is the only tuff in the East Rudolf succession whose apparent age is controversial. For example, from an extensive combination of total fusion and age spectrum $^{40}Ar/^{39}Ar$ age determinations, we estimate the best age of the Chari/Karari Tuff level at East Rudolf to be 1.32±0.05 Myr (ref. 6). This particular age was an isochron determination derived from a pumice sanidine sample which we would regard as ideal K–Ar dating material, totally free from either initial argon or argon loss discrepancies. It is instructive that identical ages have been obtained for the Chari and Karari Tuffs by conventional K–Ar dating at Berkeley (G. H. Curtis, personal communication) and similarly from Omo Tuff KNW2,

Fig. 4 Correlation diagram obtained from age spectrum analysis of KBS pumice feldspar sample FMA 274. Two distinct components seem to be present with regression line ages of a, 2.39 ± 0.01 Myr (intercept value: 297) and b, 2.42 ± 0.03 Myr (intercept value 272).
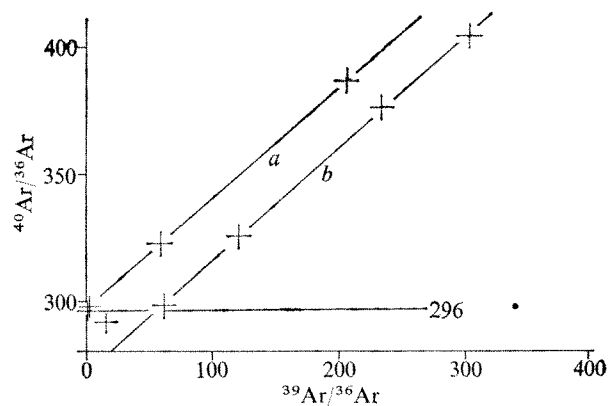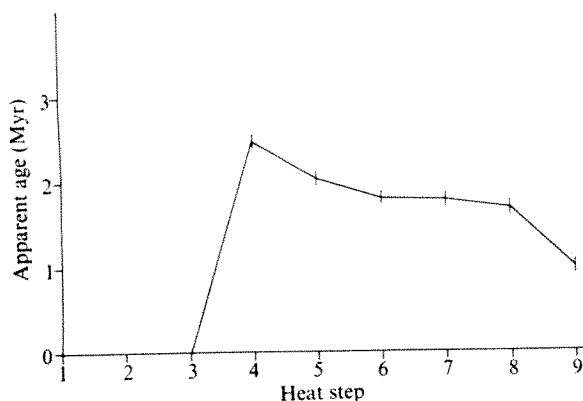
Fig. 5 $^{40}Ar/^{39}Ar$ age spectrum of KBS pumice feldspar sample FMA 517, area 131, East Rudolf (same locality as Berkeley samples ER74–131). The age spectrum has a 'peak' age of 2.47 ± 0.6 Myr then declines through a 'flat' at ~ 1.8 Myr.

which is thought by Brown[25] to be the lateral equivalent of the Chari/Karari level in southern Ethiopia. Thus, it is clear that the inconsistencies between conventional K–Ar and $^{40}Ar/^{39}Ar$ dating of the KBS Tuff most likely arise from some inherent characteristics of the samples rather than differences in the dating methods used.

Conventional K–Ar dates of feldspars from the KBS Tuff quoted by Curtis *et al.*[21] scatter between 6.90 ± 0.05 Myr and 1.50 ± 0.02 Myr. Clearly, detrital contamination was present in some of the Berkeley dating samples prepared by heavy liquid concentration alone, but, the subsequent recognition of a number of yellow; rounded, and thus probably detrital, grains in the unused portion of a dated sample does not necessarily prove that any detrital grains were present in the small aliquot of the sample actually used for the argon analysis. Thus, it is still possible that some of the apparent ages between 1.85 Myr and 2.4 Myr rejected by Curtis *et al.* may result, at least in part, from the presence of a non-detrital 2.4-Myr-old component.

Comparison

Data obtained at Cambridge from the $^{40}Ar/^{39}Ar$ age spectrum analysis of KBS pumice feldspar sample FMA 517, collected from the same locality in Area 131, East Rudolf as Berkeley samples ER74–131 are presented in Figs 5 and 6. The mean of six conventional total fusion K–Ar ages on undoubtedly detritus-free pumice feldspar and glass fractions

Fig. 6 Correlation diagram obtained from age spectrum analysis of KBS pumice feldspar sample FMA 517. Two age components seem to be present. The age of the older (*a*) cannot be defined more closely than the rather imprecise 'peak' age of 2.47 ± 0.6 Myr., but the younger (*b*) has a regression age of 1.91 ± 0.03 Myr (intercept value, 285).
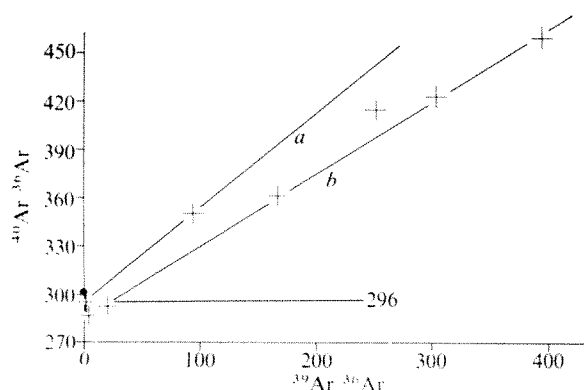
of the samples ER74-131 carried out at Berkeley was 1.82 ± 0.04 Myr. This age is regarded by Curtis *et al.* as a good estimate for the date of eruption of the KBS Tuff in Area 131, whereas, from our experience we would interpret it as being the date of total argon loss, affecting glass and feldspar crystals alike. Unfortunately, the results of the analysis of sample FMA 517 are not absolutely conclusive. The age spectrum (Fig. 5) contains a 'peak' apparent age of 2.47 ± 0.6 Myr, but at later steps the apparent age curve declines to a 'flat' apparently close to 1.8 Myr. On the correlation diagram (Fig. 6) the data fall into two groups : the younger of the two age components has an apparent age of 1.91 ± 0.03 Myr (with an intercept value of 285, suggestive of overprinting). Because insufficient data are available, the apparent age of the older component cannot be estimated with any more precision than was given to it by the peak on the age spectrum. The major feldspar component in sample FMA 517 clearly has an apparent age of 1.91 ± 0.03 Myr.

The six total fusion K–Ar ages obtained at Berkeley, the apparent age of the lower sector of the age spectrum, and a total fusion $^{40}Ar/^{39}Ar$ apparent age of 1.72 ± 0.11 Myr obtained from another subsample of pumice feldspar FMA 517 at Cambridge, must be regarded as less precise estimates of the age of this component, because they undoubtedly contain an uncorrected initial argon discrepancy and, in the case of the total fusion dates, a variable contribution from the older feldspar component.

Conclusion

By analogy with our other work on the KBS Tuff, we believe that in Area 131 the older feldspar age component represents the true age of crystallisation and the dominant younger component defines no more than the age of a widespread overprinting event, but there is nothing in the available K–Ar and $^{40}Ar/^{39}Ar$ data that makes this interpretation unique.

In these circumstances we must await confirmation from further work on samples from Area 131, especially work using methods of estimating geological time which are independent of the K–Ar dating technique. At East Rudolf, palaeomagnetic reversal chronology[26,27], and fission track dating[28,29] are the methods most likely to resolve this controversy.

One further conclusion appears inescapable: as multiple age component systems and age components with apparent initial Ar isotope ratios unlike the modern atmosphere ratio are present amongst the pumice feldspar samples from the KBS Tuff, variation in the total fusion K–Ar or $^{40}Ar/^{39}Ar$ apparent ages of samples from different localities is to be expected. Thus the conventional total fusion K–Ar dating evidence at present available cannot be used either to confirm or to deny the presence of more than one tuff or pumice population at the KBS level. It is our opinion, however, that the balance of evidence from a combination of the currently available K–Ar and $^{40}Ar/^{39}Ar$ work strongly suggests that only one tuff complex is present, it was deposited within a short interval in all areas, it contains one dominant pumice generation and this pumice is 2.42 Myr old.

1. Leakey, R. E. F., *Nature*, **226**, 223–224 (1970).
2. Behrensmeyer, A. K., *Nature*, **226**, 225–226 (1970).
3. Bowen, B. E., and Vondra, C. F., *Nature*, **242**, 391–393 (1973).
4. Fitch, F. J., and Miller, J. A., *Nature*, **226**, 226–228 (1970).
5. Fitch, F. J., Findlater, I. C., Watkins, R. T., and Miller, J. A., *Nature*, **251**, 213–214 (1974).
6. Fitch, F. J., Hooker, P. J., and Miller, J. A., in *Geological Background to Fossil Man* (edit. by Bishop, W. W.), (Geol. Soc. Lond., in the press).
7. Findlater, I. C., thesis, Univ. London (1976).
8. Turner, G., Miller, J. A., and Grasty, R. L., *Earth planet. Sci. Lett.*, **1**, 155–157 (1966).
9. Mitchell, J. A., *Geochim. cosmochim. Acta*, **32**, 781–790 (1968).
10. Dunham, K. C., Fitch, F. J., Ineson, P. R., Miller, J. A., and Mitchell, J. G., *Proc. R. Soc.*, **A307**, 251–266 (1968).
11. Fitch, F. J., Miller, J. A., and Mitchell, J. G., in *Time and Place in Orogeny* (edit. by Kent, P. E., Satterthwaite, G. E., and Spencer, A. M.), 157–196 (Geol. Soc. Lond., 1969).

[12] Brereton, R. N., *Geophys. J. R. astr. Soc.*, **27**, 449–478 (1972).
[13] Fitch, F. J., and Miller, J. A., in *Symposium on Granites, Gneisses and related rocks* (edit. by Lister, L. A.), 219–225 (Geological Society of South Africa, 1973).
[14] Fitch, F. J., Forster, S. C., and Miller, J. A., *Rep. Prog. Phys.*, **37**, 1433–1496 (1974).
[15] Fitch, F. J., and Miller, J. A., in *Earliest Man and Environments in the Rudolf Basin: Stratigraphy, Palaeoecology and Evolution* (edit. by Coppens, Y., Howell, F. C., Isaac, G. Ll., and Leakey, R. E. F.), (University of Chicago Press, Chicago, 1975).
[16] Berger, G. W., *Earth planet. Sci. Lett.*, **26**, 387–408 (1975).
[17] Fitch, F. J., Hooker, P. J., and Miller, J. A., *Geol. Mag.*, **113**, 1–10 (1976).
[18] Mellor, D. W., and Mussett, A. E., *Earth planet. Sci. Lett.*, **26**, 312–318 (1975).
[19] Curtis, G. H., in *Geological Background to Fossil Man* (edit. by Bishop, W. W.), (Geol. Soc. Lond., in the press).
[20] Curtis, G. H., *World Archaeol.*, **7**, 198–209 (1975).
[21] Curtis, G. H., Drake, R. F., Cerling, T. E., Cerling, B. L., and Hampel, J. H., *Nature*, **258**, 395–398 (1975).

[22] Maglio, V. J., *Nature*, **239**, 379–385 (1972).
[23] Cooke, H. B. S., and Maglio, V. J., in *Calibration of Hominoid Evolution*, (edit. by Bishop, W. W., and Miller, J. A.), 303–329 (Scottish Academic, Edinburgh, 1972).
[24] Brown, F. H., in *Geological Background to Fossil Man* (edit. by Bishop, W. W.), (Geol. Soc. Lond., in the press).
[25] Brown, F. H., in *Earliest Man and Environments in the Rudolf Basin: Stratigraphy, Palaeoecology and Evolution* (edit. by Coppens, Y., Howell, F. C., Isaac, G. Ll., and Leakey, R. E. F.), (University of Chicago Press, Chicago, 1975).
[26] Brock, A., and Isaac, G. Ll., *Nature*, **247**, 344–347 (1974).
[27] Brock, A., and Isaac, G. Ll., in *Earliest Man and Environments in the Rudolf Basin: Stratigraphy, Palaeoecology and Evolution* (edit. by Coppens, Y., Howell, F. C., Isaac, G. Ll., and Leakey, R. E. F.) (University of Chicago Press, Chicago, 1975).
[28] Hurford, A. J., *Nature*, **247**, 236–237 (1974).
[29] Hurford, A. J., Gleadow, A., and Naeser, C., *Nature*, **263**, 738–740 (1976).

Cloned synthetic *lac* operator DNA is biologically active

K. J. Marians & R. Wu*

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853

J. Stawinski, T. Hozumi & S. A. Narang

Division of Biological Sciences, National Research Council of Canada, Ottawa, Canada KIA OR6

A chemically synthesised duplex DNA fragment containing the sequence of the lac *operator was cloned in* E. coli *using the vehicle pMB9. Clones containing lac-pMB9 hybrid DNA produced β-galactosidase constitutively and the hybrid DNA bound the* lac *repressor specifically.*

THE nucleotide sequence of the *lac* operator has been determined by Gilbert and Maxam[1]. We have synthesised chemically an oligonucleotide duplex, 21 nucleotides long and containing the sequence of the *lac* operator, and have shown that it binds the *lac* repressor specifically[2]. Here, we report the enzymatic joining of a synthetic *lac* operator fragment with protuding 5' pA–A–T–T sequences (Fig. 1) directly to *Eco*RI endonuclease-cut pMB9 DNA[3]. After transformation, clones containing hybrid *lac*-pMB9 DNA expressed the biological activity *in vivo* of the chemically synthesised *lac* operator by producing β-galactosidase constitutively. Furthermore, the *lac* operator DNA reisolated from the hybrid DNA by cleavage with the *Eco*RI restriction enzyme[4] showed specific and strong binding to the *lac* repressor. The reisolated *lac* operator DNA with protruding 5' pA–A–T–T sequences can be joined by T4 polynucleotide ligase to produce multi-operator fragments of up to 18 copies of the *lac* operator in tandem. Some of the multiple operators were ligated with *Eco*RI endonuclease-cut pMB9 DNA, and hybrid DNA with two copies of the *lac* operator was isolated.

In addition to the specific method reported here, a more general approach for inserting any even-ended DNA fragment into plasmid DNA has been developed. In this approach[19], a chemically synthesised decanucleotide duplex containing *Bam*I restriction endonuclease[5] recognition sequence is joined to an even-ended *lac* operator DNA duplex[2] by T4 polynucleotide ligase (ref.6). Digestion of the ligated product with the *Bam*I endonuclease generated the protruding 5' pG–A–T–C sequence. This molecule was then joined to the *Bam*I endonuclease-cut pMB9 DNA. Details of these results will be reported elsewhere[19].

---

* To whom correspondence should be sent

Characterisation and analysis of hybrid plasmids

Results in Table 1 show that clones 50, 67 and 162 produced large amounts of β-galactosidase and that the DNA isolated from these clones bound the *lac* repressor extensively. The DNA from pMB9 plasmid and the three clones which produced very low levels of β-galactosidase all showed essentially no binding. The DNA from clones 67 and 162 was further examined by digesting the samples with *Eco*RI enzyme and carrying out electrophoresis using a 3%:15% polyacrylamide step gel. To detect *lac* operator-size small DNA fragments, the digest was labelled by repair synthesis[7] at the cohesive ends with DNA polymerase I and α-^{32}P-dATP. Figure 2 shows that the plasmid-size linear DNA band is clearly present in all cases. However, digests of the DNA from clones 67 and 162 also show an operator-size DNA fragment in the lower portion of the gel. The ratio of counts in the operator fragment to the plasmid DNA was nearly 1:1 for clones 67 and 162, indicating that only one copy of the *lac* operator had been inserted per plasmid. This was explored further by partially digesting the hybrid plasmids with the *Eco*RI enzyme followed by labelling, using repair synthesis (Fig. 3). Even though partial digestion products of the plasmid DNA can be seen, only one operator-size species is present in all cases. If multiple operator species were present, one would expect to see one or more bands above that of the *lac* operator DNA in the 15% gel. The small fragment was shown to be *lac* operator DNA on elution from the gels and testing its binding to the *lac* repressor. The results are shown in Table 1b. Each fragment tested gave excellent binding with the *lac* repressor. Since only the centre 21 nucleotides out of the completely repaired 31-nucleotide-long duplex had the correct sequence for the natural *lac* operator, the nucleotide sequences beyond the central 21-nucleotide sequence were not needed for the recognition by the *lac* repressor[2]. In fact, wrong sequence beyond the central 21 nucleotides did not seem to interfere with the repressor binding.

The structure of the operator fragments was confirmed by determining their nucleotide sequence. Operator fragment was labelled at the 3'-ends with ^{32}pA^{32}pA by partially repairing the protruding 5'pA–A–T–T ends using DNA polymerase I and α-^{32}P-dATP. The resulting 29-nucleotide-long strands were

separated by gel electrophoresis (Fig. 4). The sequence of each strand was then determined using a partial pancreatic DNase digestion followed by two-dimensional electrophoresis–homo-chromatography[8,9]. Figure 5a and b shows the sequences 5′ . . . A–G–C–G–G–A–T–A–A–C–A–A–T–T–G–[32]pA[32]pA

and 5′ . . . T–T–A–T–C–C–G–C–T–C–A–C–A–A–T–T–G- • [32]pA[32]pA, respectively. Written in the duplex form, these sequences overlap by nine nucleotides which confirm the expected sequence of the two 27-nucleotide-long single strands as shown in Fig. 1b.
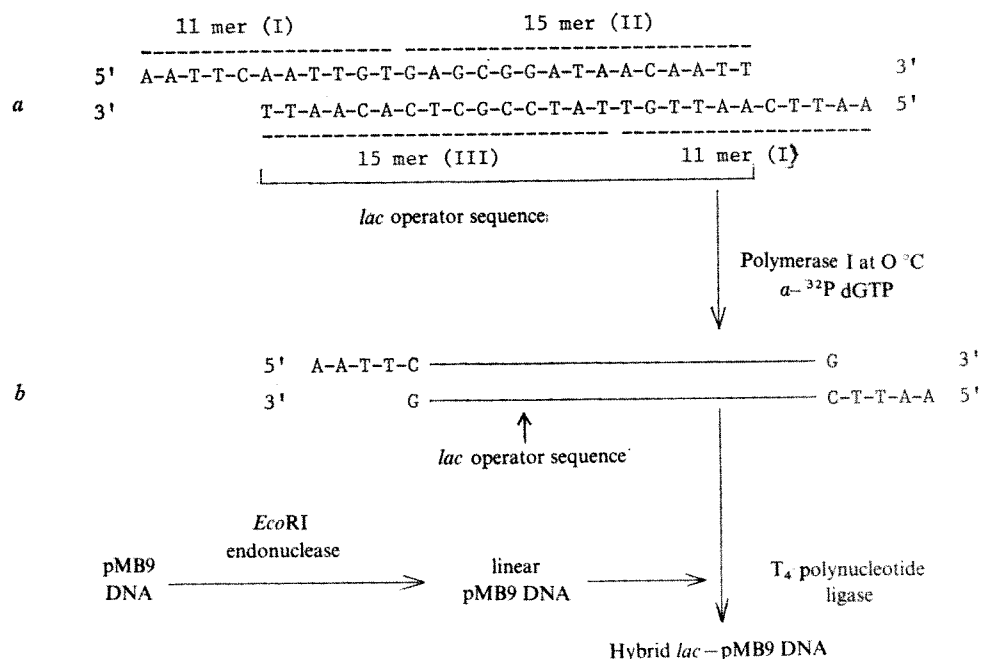
Fig. 1 Joining of the chemically synthesised *lac* operator duplex with protruding 5′ A–A–T–T sequence with *Eco*RI endonuclease-cut pMB9 DNA. The three oligonucleotides used to make the duplex fragment (structure *a*) were synthesised by the modified phosphotriester method[11–14] and are indicated by the dotted lines. Enzymatic joining of the oligonucleotides was carried out with T4 polynucleotide ligase[2] and the duplex fragment was separated from the single-stranded oligonucleotides by elution from a Sephadex G-75 column with 0.1 M NaCl. The duplex (underlined with solid line) was synthesised so that the recognition sequence of the *Eco*RI restriction endonuclease was present at both ends. Before joining the fragment with pMB9 DNA it was necessary to incorporate one dGMP residue into each of the single-stranded ends to produce structure *b*. This *lac* operator duplex with protruding 5′ A–A–T–T sequence (1 pmol) was joined to pMB9 DNA (0.3 pmol) which had been made linear by digestion with *Eco*RI endonuclease. Ligation reaction was at 12.5 °C for 20 h. The resulting hybrid *lac*–pMB9 DNA was used to transform 3 × 10⁷ competent *E. coli* HB129 cells[3] as follows: The DNA and recipient cells were mixed together and incubated at 0 °C for 30 min. The temperature of the mixture was raised to 42 °C for 2 min and then chilled to facilitate uptake of the DNA by the cells. Nine volumes of prewarmed L-broth were added and the cells allowed to recover at 37 °C for 2 h. One volume of L-broth supplemented with 10 μg ml⁻¹ of tetracycline was then added. After an additional 30 min at 37 °C the tetracycline concentration was brought up to a final level of 20 μg ml⁻¹. This cell suspension was used to inoculate 100 ml of M9 medium for the isolation of a larger amount of plasmid DNA[15]. This plasmid DNA (30 μg) which contained some hybrid *lac*–pMB9 DNA was enriched for *lac* sequences by binding it to the *lac* repressor (4.5 μg) on a Millipore filter, and eluting it with 1 ml of 1 mM isopropyl thiogalactoside (IPTG). This DNA (6% of the input), enriched for *lac* sequences, was used for a subsequent transformation on nutrient agar plates containing 20 μg ml⁻¹ of tetracycline. The frequency of transformation was 6.4 × 10⁻⁴ transformants per μg per DNA per viable cell whereas in the same conditions native pMB9 gave a frequency of 1.8 × 10⁻².

Table 1 *Lac* repressor binding

a Clone	β-Galactosidase level	Binding (%)	b Clone	Binding (%)
44	Low	0	67	28
50	High	70	162	31
67	High	72		
84	Low	4		
114	Low	3		
162	High	60		
pMB9	Low	0		

Lac repressor binding of hybrid *lac*–pMB9 plasmids and isolated operator fragments. All *lac* repressor binding assays were performed in 100 μl of binding buffer[17] containing 1.3 μg of *Hin*dII + III digested λ*plac*5 DNA (40 fmol *lac* operator), radioactive DNA sample (1–4 fmol) and 200 fmol *lac* repressor[2]. The mixture was left at room temperature for 30 min. Triplicate samples of 30 μl each were filtered through Millipore HAWP 0.45-μm filters and then washed three times with 200 μl of a washing buffer[17]. 1 ml of 1 mM IPTG was then passed through the filter and the eluate collected. Percentage binding is defined as the fraction of total input counts eluted from the filters with IPTG.

a, Assay for β-galactosidase level was carried out as follows. Plasmid pMB9 is a derivative of plasmid colicin E1 (ref. 3) and, therefore, in the unamplified state, exists in 10–20 copies per cell.

Because there are only about 10 molecules of *lac* repressor per cell, clones containing the *lac* operator ligated to plasmid pMB9 should have all *lac* repressor molecules bound to the *lac* operator on plasmids, leaving the *E. coli lac* operon derepressed, thus producing β-galacto-sidase constitutively. Clones were grown up in 1 ml L-broth containing tetracycline and were screened for β-galactosidase escape synthesis. 0.1 ml of the overnight cultures was diluted to 1 ml, and the cells were lysed with one drop each of CHCl₃ and 0.1% sodium dodecyl sulphate. *O*-nitrophenyl-β-D-galactoside (ONPG) was added to a final concentration of 0.8 mg ml⁻¹. In this qualitative assay[18], a positive indication of constitutive β-galactosidase production was the formation of a yellow colour characteristic of the action of β-galacto-sidase on ONPG. Of the 180 clones tested, three showed high levels of β-galactosidase synthesis. DNA (1 μg) from six clones (including three clones as controls) was labelled by nick translation[18] with DNA polymerase I (7 units) in 100 μl of a mixture containing 50 mM Tris-HCl (pH 7.9), 50 mM KCl, 10 mM dithiothreitol, 10 mM MgCl₂, and 8 μM of the four dNTPs (including α-^{32}P-dGTP, 15 Ci mmol⁻¹). After 2 h at room temperature, the reaction was terminated by ethanol precipitation. Input c.p.m. in the binding assay was between 500 and 1,500 c.p.m. Percentage binding had been corrected for background, determined by results without the addition of the repressor. Background was between 3–6% of the total input counts. *b*, Operator fragment eluted from a gel[7] as in Fig. 2 was found with the repressor as described earlier[17]. Input c.p.m. was between 250 and 600 c.p.m. Percentage binding as shown had been corrected for background which was 2–4%.

Fig. 2 An autoradiograph of gel electrophoresis of hybrid *lac*-plasmids digested with the *Eco*RI enzyme and labelled by repair synthesis[7]. The hybrid *lac*-pMB9 DNA was isolated after amplification by the addition of chloramphenicol to the bacterial culture[15]. The hybrid plasmid DNA (1 μg) was digested with the *Eco*RI enzyme in 20 μl of a solution containing 50 mM Tris-HCl *p*H 7.9, 100 mM NaCl, 10 mM MgCl$_2$ and 3 units of the enzyme (one unit completely digests 1 μg of DNA at 37 °C in 1 h) at 37 °C for 30 min. The reaction was terminated by chilling to 0 °C. The cohesive ends generated by the *Eco*RI enzyme were then repaired by bringing the mixture to a volume of 30 μl containing 70 mM KP$_i$(*p*H 7.0), 10 mM dithiothreitol, 2 μM α-^{32}P-dATP and dTTP (110 Ci mmol^{-1}), and 3 units of DNA polymerase I. After 2 h at 0 °C the reaction was terminated by the addition of 1 volume of 1 M Tris-HCl, followed by two volumes of ethanol. Samples were electrophoresed on a composite slab gel of 3% and 15% polyacrylamide (acrylamide–*bis*-acrylamide = 20:1) with dimensions of 20×40×0.3 cm. The running buffer was 50 mM Tris-acetate (*p*H 7.9). Xylene cyanol FF (XC) and bromophenol blue (BPB) were used as tracking dyes. Electrophoresis was at 150 V for 16 h.
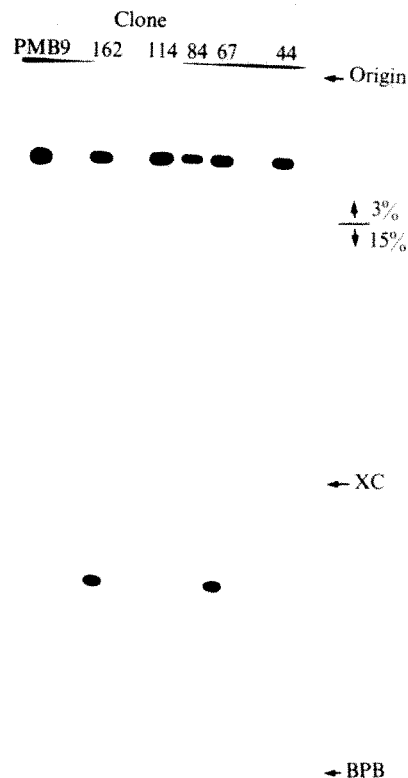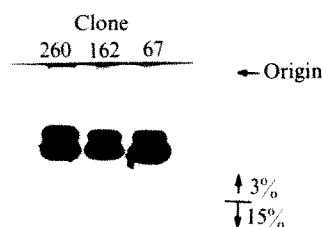
Fig. 3 An autoradiograph of gel electrophoresis of hybrid plasmids partially digested with the *Eco*RI endonuclease and labelled by repair synthesis. The procedures were the same as those given in the legend for Fig. 2, except that 1 unit of the *Eco*RI enzyme and 3 min of incubation were used.
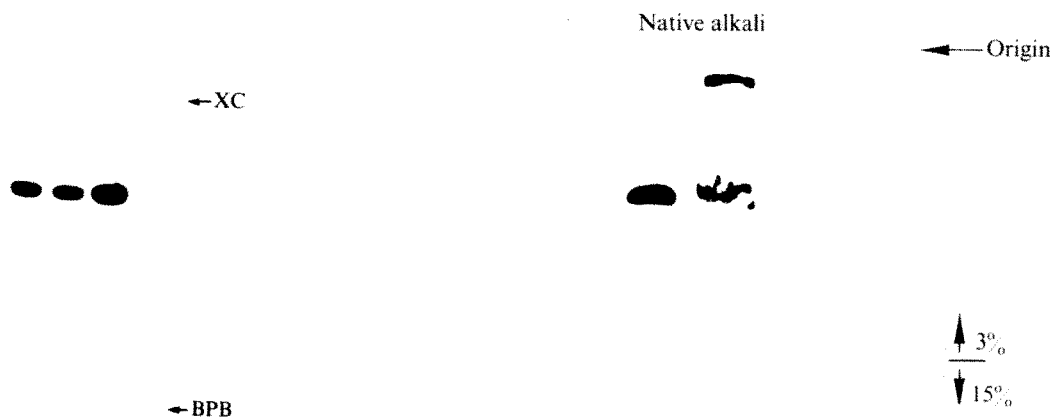
Fig. 4 Strand separation of the *lac* operator duplex. DNA from clone 67 was digested with the *Eco*RI enzyme and labelled by partial repair synthesis by DNA polymerase I using only α-^{32}P-dATP. This now yields DNA fragments with single-strand length of 29 nucleotides. After ethanol precipitation one half the material was dissolved in standard loading buffer and the other half was dissolved in 0.3 N NaOH–10% glycerol. Bromophenol blue and xylene cyanol FF were added to the alkali sample just before loading on a 20×40×0.3-cm 3%:15% step gel of polyacrylmide (acrylamide–*bis*-acrylamide = 20:1). The gel was run in Tris-borate buffer (*p*H 8.3) at 400 V. Two distinct bands corresponding to the separate strands of the *lac* operator fragment can be seen in the slot containing the alkali sample. This can be compared with the mobility of the native fragment which is in the other slot. These bands were eluted and used for sequence analysis as described in the text.

Fig. 5 *a* and *b*, Autoradiographs of two-dimensional electrophoresis–homochromatography of partial DNase digests of the separated strands of the *lac* operator fragment reisolated from hybrid plasmid DNA as in the legend for Fig. 2 and 4. The sequences were deducted by the quantitative mobility-shift analysis[9]. The sequence of 17 nucleotides from one strand and 19 nucleotides from the other strand overlaps by 9 nucleotides to give the sequence of the two 29-nucleotide-long single strands which confirms that 27-nucleotide long sequence as shown in Fig. 1*b*
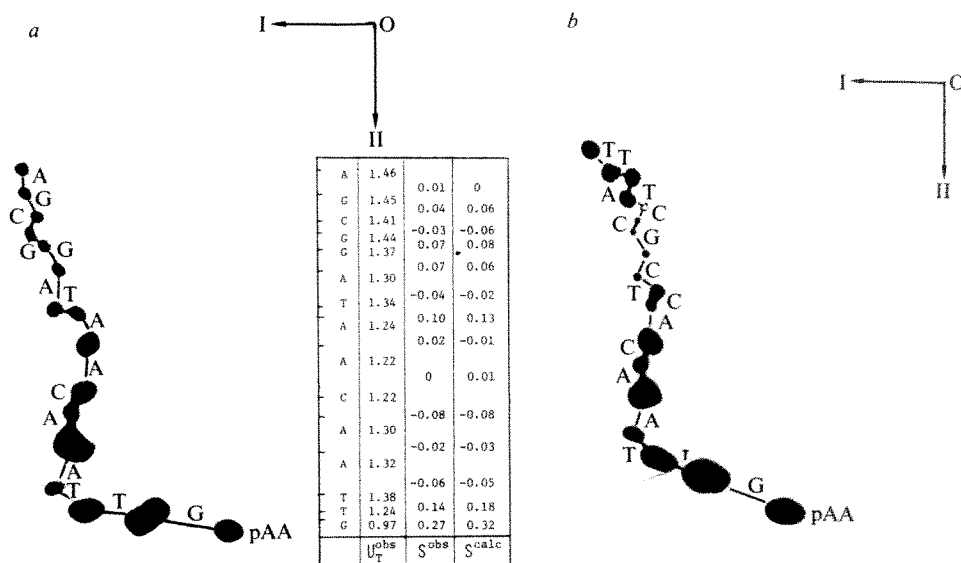
Fig. 6 An autoradiograph of gel electrophoresis of polymeric *lac* operators. Cloned *lac* operator fragment reisolated from hybrid *lac*–pMB9 DNA by *Eco*RI enzyme digestion was labelled at the 5′ end with ³²P and polynucleotide kinase[9]. The fragment was eluted from a gel and then joined end-to-end with itself (*lac* operator DNA concentration was 130 ng ml⁻¹, using T4 polynucleotide ligase at 12.5 °C. The reaction mixture was analysed on a 10% polyacrylamide gel with a 3% stacking gel, using 40 mM Tris-borate (*p*H 8.3) electrophoresis buffer. When no ligase was used (result not shown), only the monomeric operator (band 1) was present. The lighter bands in between dark bands are presumably circular oligomeric and polymeric *lac* operators.
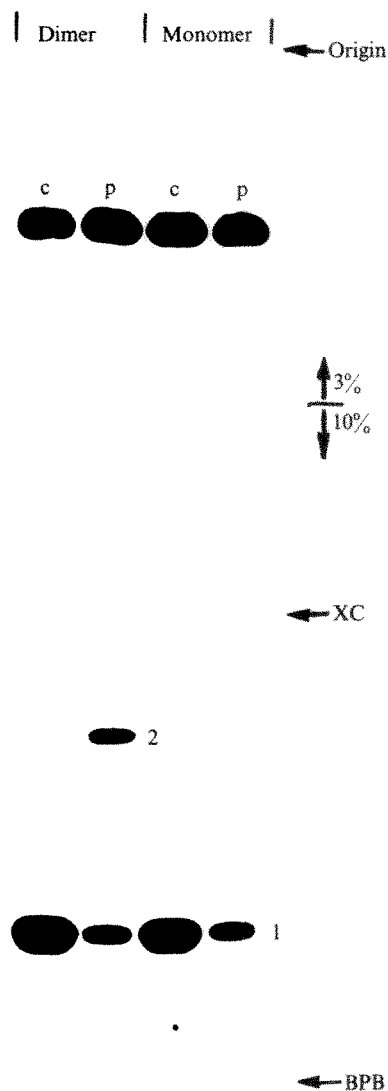
Fig. 7 Characterisation of a plasmid containing two copies of the *lac* operator in tandem. The band corresponding to a dimer of the *lac* operator fragment was eluted from the gel shown in Fig. 6. This material was then ligated with *Eco*RI endonuclease-cut pMB9 DNA and used to transform HB129 cells. β-galactosidase constitutive clones were identified by their blue colour on X-gal plates[16]. The number of copies of *lac* operator DNA contained in any particular clone was determined by (1) partially digesting (lane *p*) the hybrid plasmid DNA with *Eco*RI endonuclease, and (2) by complete digestion (lane *c*). The digest was then labelled by repair synthesis using α-³²P-dATP and DNA polymerase I. The results were analysed by electrophoresis on a 20 × 40 × 0.3 cm 3%:10% step gel of polyacrylamide (acrylamide–*bis*-acrylamide, 20:1) run in Tris-borate buffer. DNA from clone 67 (marked monomer) are shown for comparative purposes. For experiments with complete *Eco*RI endonuclease digestion, the radioactive DNA bands were cut out from the polyacrylamide gel and counted for ³²P. In the dimeric hybrid, *lac*₂–pMB9 DNA (marked dimer), the ratio of counts in the operator fragment (band 1 in the 10% gel) to the plasmid DNA (the dark band above the 3%:10% junction) was close to 2:1 (60,000 c.p.m:29,500 c.p.m). In the monomeric hybrid (marked monomer) the ratio of counts was 1:1.

Since the yield of *lac* operator from the *lac*–pMB9 hybrid is relatively low, it will be necessary to produce a plasmid with multiple copies of the *lac* operator. To this end, we attempted to determine whether the cloned *lac* operator fragment can be ligated to itself to form polymeric *lac* operators before ligation to the plasmid DNA. Figure 6 shows that one can clearly distinguish on this polyacrylamide gel multi-operator DNA bands which contain up to 18 copies of the *lac* operator (joined together by the protruding sequence 5'pA–A–T–T). These multi-operator DNA molecules appeared to be linear instead of circular in structure as reflected by the percentage of ^{32}P counts released by treating each DNA band (up to the nonamer) with bacterial alkaline phosphatase. A dimeric *lac* operator fragment has been inserted into plasmid pMB9 (Fig. 7). When the hybrid *lac*$_2$–pMB9 DNA was partially digested with *Eco*RI endonuclease a monomeric (band 1) and dimeric (band 2) *lac* operator were found. Insertion of oligo-operator DNA fragments into the plasmid and the characterisation of these hybrid molecules are in progress.

Potential of this technique

The technique described here provides a way of producing large amounts of homogeneous and biologically active *lac* operator DNA for physicochemical studies. The first method (Fig. 1) can also be used to insert any similarly constructed synthetic DNA into a cloning vehicle. The second method[19] which involves end-to-end joining of an even-ended DNA to a synthetic DNA duplex containing the *Bam*I (or other) restriction endonuclease recognition sequence can be used to insert any even-ended DNA molecules into a cloning vehicle. Our technique serves as an example of a general approach in introducing a *lac* operator in front of a gene or a group of genes, so that the gene can be turned on or off by *lac* repressor with or without IPTG.

1 Gilbert, W., and Maxam, A., *Proc. natn. Acad. Sci. U.S.A.*, 70, 3581–3584 (1973).
2 Bahl, C. P., Wu, R., Itakuru, N., Katagiri, N., and Narang, S. A., *Proc. natn. Acad. Sci. U.S.A.*, 73, 91–94 (1976).
3 Rodriquez, R. H., Bolivar, F., Goodman, H. M., Boyer, H. W., and Betloch, M., *Proceedings of the 1976 ICN-UCLA Symposium on Molecular Biology* (Academic, in the press).
4 Hedgpeth, J., Goodman, H. M., and Boyer, H. W., *Proc. natn. Acad. Sci., U.S.A.*, 69, 3448–3452 (1972).
5 Wilson, G. A., and Young, F. E., *J. molec. Biol.*, 97, 123–125 (1975).
6 Sgaramella, V., van de Sande, J. H., and Khorana, H. G., *Proc. natn. Acad. Sci. U.S.A.*, 67, 1468–1475 (1970).
7 Wu, R., Jay, E., and Roychoudhury, R., in *Methods in Cancer Research*, 12, 87–176 (1976).
8 Brownlee, G. G., and Sanger, F., *Eur. J. Biochem.*, 11, 395–399 (1969).
9 Tu, C. D., Jay, E., Bahl, C. P., and Wu, R., *Analyt. Biochem.*, 73 (in the press).
10 Heyneker, H., *et al.*, *Nature*, 263, 748–752 (1976).
11 Itakura, K., Bahl, C. P., Katagiri, N., Michniewicz, J. J., Wightman, R. H., and Narang, S. A., *Can. J. Chem.*, 51, 3649–3651 (1973).
12 Itakura, K., Katagiri, N., Bahl, C. P., Wightman, R. H., and Narang, S. A., *J. Am. chem. Soc.*, 97, 7327–7332 (1975).
13 Katagiri, N., Itakura, K., and Narang, S. A., *J. Am. chem. Soc.*, 97, 7332–7336 (1975).
14 Itakura, K., Katagiri, N., Narang, S. A., Bahl, C. P., Marians, J. K., and Wu, R., *J. biol. Chem.*, 250, 4592–4600 (1975).
15 Katz, L., and Helinski, D. R., *J. Bact.*, 119, 450–460 (1973).
16 Miller, J. H., *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, 1972).
17 Riggs, A. D., and Bourgeois, S., *J. molec. Biol.*, 34, 361–364 (1968).
18 Maniatas, T., Jeffrey, A., and Kleid, D., *Proc. natn. Acad. Sci. U.S.A.*, 72, 1184–1188 (1975).
19 Bahl, C. P., Marians, K. J., Wu, R., Stawinsky, J., and Narang, S. A., *Gene*, 1 (in the press).

Synthetic *lac* operator DNA is functional *in vivo*

Herbert L. Heyneker, John Shine, Howard M. Goodman & Herbert W. Boyer

University of California, San Francisco, California 94143

John Rosenberg & Richard E. Dickerson

California Institute of Technology, Pasadena, California 91106

Saran A. Narang

National Research Council of Canada, Ottawa, Ontario K1A OR6, Canada

Keiichi Itakura, Syr-yaung Lin & Arthur D. Riggs

City of Hope National Medical Center, Duarte, California 91010

Blunt end ligation with an EcoRI *octanucleotide "linker" was used to construct a plasmid containing chemically synthesised* lac *operator DNA. This plasmid renders the host bacterium constitutive for β-galactosidase.*

THE specific binding of the *lac* repressor protein to the *lac* operator, a defined nucleotide sequence in the DNA of *Escherichia coli*, controls the expression of the nearby structural genes of the *lac* operon[1]. The nucleotide sequence in the region of the *lac* operator is known[2,3] and recently the *lac* operator (21 nucleotide base pairs, see Fig. 1) has been synthesised by a phosphotriester chemical method[4,5]. This chemically synthesised *lac* operator DNA functions *in vitro*, as evidenced by repressor binding activity[5,6].

To examine the *in vivo* properties of a chemically synthesised genetic element and to have a convenient con-tinuous source of a small *lac* operator DNA, the chemically synthesised *lac* operator was cloned in a bacterial plasmid. The scheme which produced an easily clonable operator DNA fragment is depicted in Fig. 1. Using blunt end ligation by T4 DNA ligase[7], a chemically synthesised octadeoxyribonucleotide containing the *Eco*RI endonuclease recognition sequence[8] was joined to the 21-nucleotide *lac* operator sequence. The resulting RI–operator DNA fragment was then inserted into the plasmid pMB9 (ref. 9) by using the *Eco*RI cohesive termini and established procedures[10]. The pMB9 plasmid was used because it has one *Eco*RI restriction site, carries a tetracycline resistance gene, and can be amplified by chloramphenicol to yield several thousand copies per cell[9,11].

The covalently joined molecules of *lac* operator and *Eco*RI substrate which we refer to as the RI–operator fragment are totally chemically synthesised DNA molecules. The synthesis of β-galactosidase is constitutive in cells

RNA polymerase

Repressor–protected DNA

Met Thr Ile Thr Asp

Ribosome–protected RNA

z mRNA

```
CTTTATGCTTCCGGCTCGTATGT TGTGTG AATTGTGAGCGGATAACAATT TCACACA GGAAACAGCTATGACCATGATTACGGATTCACTGG
GAAATACGAAGGCCGAGCATACA ACACAC TTAACACTCGCCTATTGTTAA AGTGTGT CCTTTGTCGATACTGGTACTAATGCCTAAGTGACC
```

-30   -20   -10   0   10   20   30   40   50   60

```
TGAATTCA                AATTGTGAGCGGATAACAATT             TGAATTCA
          +                                        +
ACTTAAGT                TTAACACTCGCCTATTGTTAA             ACTTAAGT
```

T4 DNA ligase

```
TGAATTCA AATTGT GAGCGGATA ACAATT TGAATTCA
ACTTAAGT TTAACA CTCGCCTAT TGTTAA ACTTAAGT
```

EcoRI endonuclease

```
AATTCA AATTGT GAGCGGATA ACAATTTG
GT TTAACA CTCGCCTAT TGTTAA ACTTAA
```

Fig. 1 Experimental plan for cloning of the *lac* operator. The nucleotide sequence at the top of the figure represents most of the control region of the *lac* operon. The 5′ termini are on the left of the upper strands as depicted here. The sites for RNA polymerase binding, mRNA initiation, and so on are depicted[3]. The *lac* operator, where repressor binds, is easily identified by the boxed regions of symmetry. The central operator, where all operator-constitutive mutations occur, is from nucleotide 0 to +20. The extended operator noted by Dickson *et al.*[3], is from −7 to +27. The second line shows the nucleotide sequences of the chemically synthesised DNAs. The 21-nucleotide central operator was synthesised by a phosphotriester method[4,5,18] and the octanucleotide *Eco*RI substrate was made by phosphodiester synthesis[8]. The chemically synthesised DNAs were phosphorylated by polynucleotide kinase, joined by T4 DNA ligase and treated with *Eco*RI endonuclease to produce cohesive ends. The *Eco*RI endonuclease-treated fragment was cloned in the *Eco*RI site of the plasmid pMB9 cloning vehicle[9]. Additional details are given in the legend of Fig. 2.

carrying the RI–operator–pMB9 plasmid because there are about 30 copies of this plasmid per cell in normal growth conditions[9]. The extra operators titrate the 10 to 20 repressor molecules per cell[12], so the single *lac* operon in the host chromosome is not repressed. Therefore, the experiments reported here demonstrate that the genetic constitution of a cell can be altered by the incorporation of chemically synthesised DNA.

The technique devised for cloning the *lac* operator DNA utilises the chemically synthesised octadeoxyribonucleotide containing the substrate for the *Eco*RI endonuclease as a "linker". The addition of short restriction site "linkers" to DNA of any source provides novel flexibility to cloning technology.

Blunt end ligation

Following the procedure given in the legend of Fig. 1, about a 20-fold molar excess of unlabelled *Eco*RI octadeoxyribonucleotide was added to the 5′ ^{32}P-labelled *lac* operator DNA and treated with T4 DNA ligase at 12.5 °C, a temperature slightly below the T_m of the *Eco*RI octanucleotide duplex structure[8]. The *Eco*RI linker was joined to the 21 base pair operator DNA by the blunt-end DNA joining activity of T4 DNA ligase, a property of this enzyme discovered by Sgaramella and Khorana[7] and recently used by W. Gilbert (personal communication) for plasmid construction.

About 50% of the 5′-terminal phosphates on the *lac*

operator DNA were resistant to alkaline phosphatase after a 3-h incubation. A sample of this reaction, along with separate samples of terminally-labelled *lac* operator and *Eco*RI substrate treated separately with T4 DNA ligase, were subjected to electrophoresis on a 12% polyacrylamide gel. An autoradiogram of the gel reveals that the *lac* operator DNA and the *Eco*RI DNA were joined (Fig. 2); the ligated samples of the *Eco*RI linker and the combined *Eco*RI linker plus *lac* operator had molecules of higher molecular weight which must have been formed by the action of the T4 DNA ligase. *Eco*RI endonuclease digestion eliminated the larger molecules, and a new oligonucleotide, 29 nucleotide base pairs in length, appeared.

Cloning

The RI–operator DNA produced by *Eco*RI endonuclease digestion was ligated to *Eco*RI endonuclease-treated pMB9 plasmid DNA and used to transform *E. coli* RR1, which has an i^+z^+ genotype[9]. After 3 h growth for expression of tetracycline resistance, samples of the culture were plated on minimal medium containing 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal) and tetracycline. X-gal is not an inducer for the *lac* operon but is a substrate for β-galactosidase[13,14]. After 48 h incubation at 37 °C, 0.3% of the tetracycline-resistant colonies were blue, indicating that these colonies were constitutively synthesising β-galactosidase, and releasing the blue dye 5-bromo-4-chloro-indigo. Since there are about 30 copies of the pMB9 plasmid per

cell and only about 10 *lac* repressor molecules, synthesis of β-galactosidase becomes constitutive if the plasmid contains a functional *lac* operator. Four independent blue colonies were chosen for further studies (plasmid numbers pHH1–pHH4). About a 40-fold increase in the uninduced level of β-galactosidase was observed for each strain.

In vitro characterisation and sequence analysis

Plasmid DNA was prepared from the transformants carrying the independently derived RI–operator plasmids. After *Eco*RI endonuclease digestion, the DNA was treated with alkaline phosphatase, extracted with phenol, and the 5′ terminal labelled with [32]P. The labelled DNA was subjected to electrophoresis on a 20% polyacrylamide gel and the ^{32}P located by autoradiography. Two bands were observed for all four plasmids. These were excised and found to have equal radioactivity, indicating that there is one RI–operator per plasmid. The fastest migrating component was estimated to be 29 nucleotides long, the expected length of the RI–operator fragment (Fig. 2). The short RI–operator fragment was preparatively separated from the 3.5×10^6 dalton plasmid DNA by filtration on Agarose A-50 followed by concentration on hydroxylapatite, or by differential elution from DEAE–cellulose (legend of Fig. 5).

Purified RI–operator DNA from pHH2 was labelled at the 5′ end, alkali denatured and the two short complementary strands separated by electrophoresis on an 8% polyacrylamide gel[15]. The nucleotide sequence was determined by a procedure developed by Maxam and Gilbert (personal communication). The method depends on base specific modification and cleavage of the DNA followed by separation according to length on a 20% acrylamide gel in 7 M urea. The sequence can be read from the autoradiogram shown in Fig. 3. The nucleotide sequence was identical to the 21-nucleotide sequence of the chemically synthesised *lac* operator with an AATTCA sequence at the 5′ end and a TG sequence at the 3′ end (see Fig. 1). The nucleotide sequences of the RI–operator DNA from all four plasmids were compared without separation of the complementary strands and no differences were evident (data not shown).

Repressor binding

The RI–operator in uncleaved plasmid DNA binds repressor very well (Fig. 4). Because supercoils in the plasmid DNA might alter the affinity of *lac* operator for *lac* repressor[16], we converted the pHH2 plasmid DNA to a linear form with *Hin*dIII endonuclease. This enzyme cuts the pMB9 plasmid

Fig. 2 Blunt end ligations and cloning. The chemically synthesised *lac* operator DNA was labelled at the 5′ termini by treatment with T4 polynucleotide kinase (2–5 units per 100 pmol of 5′ ends) in a reaction containing 50 mM Tris-HCl, *p*H 9.5, 10 mM MgCl$_2$, 5 mM dithiothreitol, 5% glycerol and 2 μM γ-^{32}P-ATP (1,000 Ci mmol^{-1}). After 1 h incubation at 37 °C, additional T4 kinase and ATP (to 100 μM final concentration) were added to the reaction and incubation was continued for an additional 30 min to obtain near quantitative phosphorylation of the DNA. The labelled DNA was separated from residual ATP by chromatography on Sephadex G-50. 75 ng of ^{32}P-labelled 21-nucleotide *lac* operator DNA was joined with 550 ng of unlabelled *Eco*RI linker DNA by treatment with 0.4 units of T4 DNA ligase[19] in a 10-μl reaction volume (20 mM Tris-HCl, *p*H 7.6, 10 mM MgCl$_2$, 10 mM dithiothreitol, and 0.4 mM ATP). The reaction was incubated for 3 h at 12.5 °C. For cloning, a sample of the ligated mixture of *Eco*RI linker DNA and operator DNA (about 2.5 pmol of operator) was combined with 0.2 pmol of pMB9 DNA and treated with 50 units of *Eco*RI endonuclease[8] at 37 °C for 60 min in 30 μl (final volume) of 100 mM Tris-HCl, *p*H 7.5, 10 mM MgCl$_2$, 100 mM NaCl, and 0.02% NP40. After inactivation of the *Eco*RI endonuclease by chloroform extraction, 5 μl of 10-times concentrated ligase buffer (0.66 M Tris-HCl, *p*H 7.6, 66 mM MgCl$_2$, 0.1 M dithiothreitol, and 4 mM ATP) and 0.4 units of T4 DNA ligase were added and the volume adjusted to 50 μl. The reaction was incubated for

16 h at 12 °C, dialysed against 30 mM CaCl$_2$, and used to transform strain RR1[9] according to an established procedure[10]. The transformed culture was plated on glucose minimal medium[11] containing 40 μg ml^{-1} of 5-bromo-4-chloro-3-indolyl-β-D-galactoside and 10 μg ml^{-1} of tetracycline. There were approximately 5,000 tetracycline resistant transformants per ml of culture. The reaction was followed by electrophoresis of the products in a 12% polyacrylamide gel containing 7 M urea, 50 mM Trisborate *p*H 8.3, 1 mM EDTA. An autoradiogram (Kodak NS X-ray film) of the gel is shown. Migration of the DNAs is from top to bottom. Lane 1: ^{32}P-labelled *lac* operator DNA. The slowest migrating band contains the intact operator DNA strands and the faster migrating components are presumably incomplete reactants or other contaminants. Lane 2: Ligated mixture of unlabelled *Eco*RI-linkers and ^{32}P-labelled *lac* operator DNA. Distinct oligonucleotide bands larger than the 21-nucleotide *lac* operator are evident and represent the operator DNA to which one, two, three or more *Eco*RI-linkers have been ligated. Lane 3: *Eco*RI digested ligation mixture. Treatment of the material shown in lane 2 with *Eco*RI restriction endonuclease reduces the large molecules to a size equal to or smaller than the RI–operator shown in lane 4. Thus the majority of the ligated material seen in lane 2 represents operator DNA bounded by *Eco*RI-linkers rather than multimers of the *lac* operator DNA. Lane 4: RI–operator DNA purified from the plasmid pHH2. The purification and ^{32}P-labelling of the RI–operator are described in the legend to Fig. 3.

Fig. 3 DNA sequence analysis of the RI–operator fragment. Purified pHH2 plasmid DNA was treated with *Eco*RI endonuclease to release the small operator fragment. The cleaved DNA was dephosphorylated by treatment with alkaline phosphatase (1 unit per 200 pmol of 5′ termini in 25 mM Tris-HCl, *p*H 8.0, 1 mM EDTA), followed by phenol extraction and G-50 Sephadex gel filtration. The dephosphorylated DAN was ³²P-labelled as described in the legend of Fig. 2. The small RI–operator DNA was separated from the 3.5 megadalton linear plasmid DNA by chromatography on DEAE–cellulose. The RI–operator DNA was eluted with 0.8 M triethylammonium carbonate at *p*H 8.0 and concentrated by lyophylisation. The complementary strands of the RI–operator were separated by denaturation in 0.1 M NaOH followed by electrophoresis on an 8% polyacrylamide gel in 50 mM Tris-borate *p*H 8.3, and 1 mM EDTA[15]. The two strands separated completely and, after elution from the gel, were used for sequence analysis. The sequence determinations were carried out according to a protocol kindly supplied by Maxam and Gilbert (to be published). The procedure depends on base specific cleavage of 5′-labelled DNA, followed by separation of the fragments according to size by electrophoresis on 20% acrylamide gels under denaturing conditions (7 M urea). Preferential cleavage at guanosine residues was achieved by methylation with dimethylsulphate, depurinisation by heating to 90 °C at *p*H 7.0, followed by alkali treatment. Cleavage at adenosine residues was achieved by depurinating the methylated DNA at 0 °C for 1 h in 0.1 N HCl. Preferential modification and subsequent cleavage at cytidine residues was obtained by treatment with hydrazine in the presence of 1 M NaCl. Cleavage at thymidine residues results when hydrazine treatment is carried out in the absence of NaCl. The darkest band at each fragment size indicates which condition caused preferential cleavage, and thus identifies the base at which cleavage occurred to give the fragment. In practice, each cleavage condition is only partially specific, but the cleavage pattern for all four conditions run in parallel allows the sequence to be determined, especially when both strands are analysed. The right four lanes of the autoradiogram shows the cleavage pattern for the upper strand. The left four lanes show the lower strand. For each set the G-specific cleavage is on the left followed by A-specific, T-specific, and C-specific. The first two nucleotides at the 5′ end of each strand were not determined by this method but are known from the specificity of *Eco*RI endonuclease.
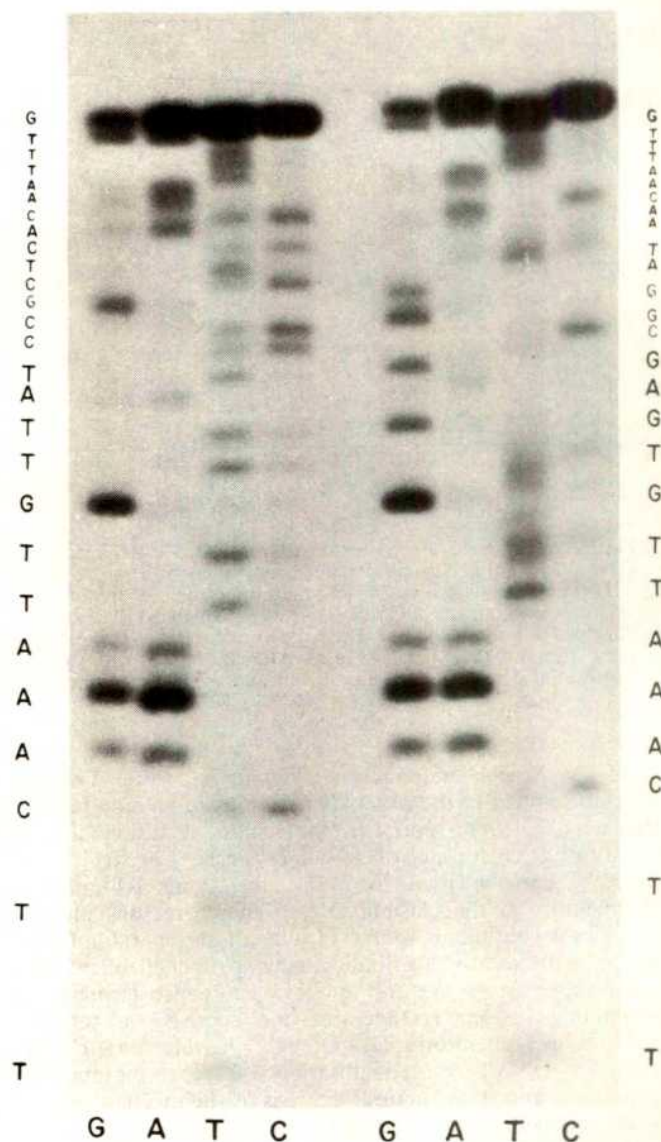
about 200 base pairs from where the RI–operator is located[9]. Linear pHH2 DNA competes even more effectively for repressor; only a threefold molar excess gives 50% competition with ³H-λ*plac* DNA. This result strongly suggests that when the RI–operator is sequestered in high molecular weight DNA, it has an affinity for repressor close to that of natural *lac* operator.

Direct filter binding experiments with purified ³²P-labelled RI–operator DNA fragments were also done and they demonstrate clearly that the small RI–operator DNA fragment binds *lac* repressor (Fig. 5). The filter retention is eliminated by isopropylthiogalactoside (IPIG), an inducer of the *lac* operon. The plateau of filter retention is only 20% for the RI–operator DNA, whereas the filter retention

of λ*plac* DNA is usually about 50%. The lower retention efficiency for the RI–operator DNA probably is a reflection of the lower binding affinity of the small fragment for repressor. Our preliminary competition data indicate that the excised RI–operator fragment has at least a 100-fold reduced affinity for repressor, compared with intact pHH2 plasmid DNA.

Chemically synthesised DNA and cloning

In this paper and the paper by Marians *et al.*[17], are reported the first successful construction of plasmids carrying a functional genetic element derived from chemically synthesised DNA. The plasmids described here carry the 21-nucleotide sequence of the *lac* operator bounded on both

Fig. 4 Competition for *lac* repressor by plasmid DNA. Various amounts of unlabelled plasmid DNA were mixed with 0.025 µg of ³H-λ*plac* DNA and then an equimolar amount of *lac* repressor was added. This amount of repressor was found to give 85% saturation of the operator in ³H-λ*plac* DNA in the absence of competing DNA. The reaction volume was 0.4 ml and the buffer was 10 mM KCl, 10 mM Tris-HCl (*p*H 7.4), 0.1 mM EDTA, 0.1 mM dithiothreitol, 5% (v/v) dimethyl-sulphoxide, and 50 µg ml⁻¹ of bovine serum albumin. After 30 min incubation at room temperature, 0.15-ml aliquots were filtered in duplicate through nitrocellulose filters according to the procedure of Riggs *et al.*[20], and washed once with 0.2 ml of buffer without bovine serum albumin or dithiothreitol. ●, pMB9 plasmid DNA; ○, pHH2 plasmid DNA; △, *Hind*III endonuclease-treated pHH2 plasmid DNA.
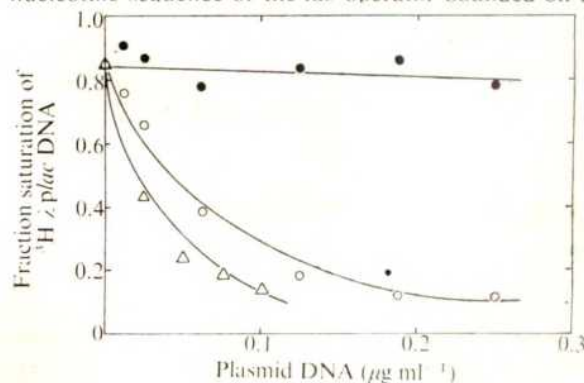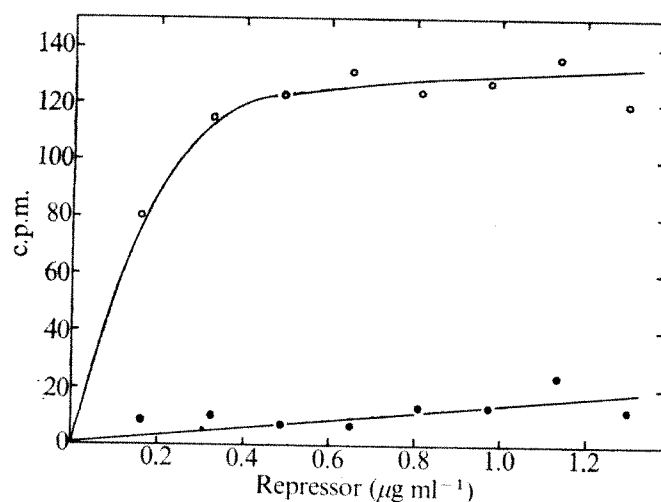
Fig. 5 Binding of *lac* repressor to purified ³²P-RI–operator DNA. The purified small RI–operator fragment was labelled with ³²P-phosphate by T4 polynucleotide kinase treatment. Varying amounts of *lac* repressor were added to 2.6 ng of purified RI–operator DNA. The reaction volume was 0.5 ml and the buffer was that used in Fig. 4. After 20 min at room temperature, 0.2 ml aliquots were filtered and washed as in Fig. 4. ●, No IPTG; ○, 10⁻³ M IPTG.

find that the RI–operator sequence, when it is sequestered in large molecular weight DNA, has an affinity for repressor within a factor of three of that of natural *lac* operator in λ*plac* DNA. Therefore, the symmetrical sequences in the "extended operator" are not obviously involved in this protein–DNA interaction. When RI–operator is cut from the plasmid DNA, the affinity for repressor is lowered greatly. Therefore, it seems necessary to have DNA on both sides of the 21-base-pair operator for "normal" interactions, but the sequence of this DNA is probably not critical.

The method used to generate the RI–operator fragment, that of the blunt end joining of a short restriction enzyme substrate (linker) to each end of the DNA of interest, is generally useful for DNA manipulation and cloning. For example, *Eco*RI "linkers" or other endonuclease substrate linkers now can be added to any blunt ended DNA molecule. The combination of chemical synthesis technology with cloning technology will provide a powerful approach to many problems, including the chemical synthesis of long DNA sequences.

This work was supported by grants from the NSF, the NIH, and the National Research Council of Canada. H.L.H. is the recipient of a fellowship grant from the Netherlands Organization for the Advancement of Pure Research (ZWO). J.S. is the recipient of a CSIRO postdoctoral fellowship.

sides by *Eco*RI endonuclease restriction sites. The *lac* operator cloned in the plasmids functions *in vivo* by binding the available *lac* repressor molecules, resulting in the constitutive synthesis of β-galactosidase. *Eco*RI endonuclease can be used to excise the small RI–operator fragment from the plasmid DNA. Therefore, this plasmid provides a continuous source of a small *lac* operator DNA, which will greatly facilitate many physicochemical and other experiments. We are now trying to construct plasmids containing several RI–operators polymerised in tandem. Such a polyoperator plasmid will provide enough *lac* operator DNA for crystallisation of repressor–operator complexes and the structural analysis of the mechanism of a sequence-specific protein–DNA interaction.

These studies show that the 21-nucleotide sequence of the "central" lac operator provides for almost all of the specific *in vitro* interactions with *lac* repressor. As shown in Fig. 1, Dickson *et al.*[3] noted additional regions of symmetry on each side of the central 21-base-pair operator. This observation led to speculation that the repressor might interact specifically with a larger DNA sequence (35 base pairs). We

1 Jacob, F., and Monod, J., *J. molec. Biol.*, **3**, 318–356 (1961).
2 Gilbert, W., and Maxam, A., *Proc. natn. Acad. Sci. U.S.A.*, **70**, 3581–3584 (1973).
3 Dickson, R. C., Abelson, J., Barnes, W. M., and Reznikoff, W. S., *Science*, **187**, 27–35 (1975).
4 Itakura, K., *et al.*, *J. biol. Chem.*, **250**, 4592–4600 (1975).
5 Bahl, C. P., Wu, R., Itakura, K., Katagiri, N., and Narang, S. A., *Proc. natn. Acad. Sci. U.S.A.*, **73**, 91–94 (1976).
6 Lin, S., *et al.*, *ICN-UCLA Symposium on Genetic Regulatory Mechanisms* (Academic, New York, in the press).
7 Sgaramella, V., Van de Sande, J. H., and Khorana, H. G., *Proc. natn. Acad. Sci. U.S.A.*, **67**, 1468–1475 (1970).
8 Greene, P. J., *et al.*, *J. molec. Biol.*, **99**, 237–261 (1975).
9 Rodriguez, R. L., Bolivar, F., Goodman, H. M., Boyer, H. W., and Betlach, M., *ICN-UCLA Symposium on Genetic Regulatory Mechanisms* (Academic, New York, in the press).
10 Chang, A. C. Y., and Cohen, S. N., *Proc. natn. Acad. Sci. U.S.A.*, **71**, 1030–1034 (1974).
11 Clewell, D. B., and Helinski, D. R., *Proc. natn. Acad. Sci. U.S.A.*, **62**, 1159–1166 (1969).
12 Gilbert, W., and Müller-Hill, B., *Proc. natn. Acad. Sci. U.S.A.*, **56**, 1891–1898 (1966).
13 Scaife, J., and Beckwith, J. R., *Cold Spring Harb. Symp. quant. Biol.*, **31**, 403 (1966).
14 Miller, J. H., *Exp. molec. Genet.* (Cold Spring Harbor, New York, 1972).
15 Hayward, G. S., *Virology*, **49**, 342–344 (1972).
16 Wang, J. C., Barkley, M. D., and Bourgeois, S., *Nature*, **251**, 247–249 (1974).
17 Marians, K., Wu, R., Stawinski, J., Hozumi, T., and Narang. S. A., *Nature*, **263**, 744–748 (1976).
18 Itakura, K., Bahl, C. P., Katagiri, N., and Narang, S. A., *Can. J. Chem.*, **51**, 3649 (1973).
19 Panet, A., *et al.*, *Biochemistry*, **12**, 5045–5050 (1973).
20 Riggs, A. D., Suzuki, H., and Bourgeois, S., *J. molec. Biol.*, **48**, 67–83 (1970).

letters to nature

The hard X-ray spectrum of Cyg X-1 during the transition in November 1975

We wish to report an observation of the hard X-ray spectrum of Cyg X-1 during a transition to the high state in November 1975. The measurement was obtained with a balloon-borne X-ray detector in the 25–150 keV range. This third upward transition of Cyg X-1 in 1975 was also detected in the few-keV energy range by the Ariel V all-sky monitor (S. S. Holt, L. J. Kaluzienski, E. A. Boldt, and P. J. Serlemitsos, unpublished). Our high energy data seem to confirm the characteristic spectral time variation of Cyg X-1: a soft X-ray intensity

increase is correlated with a decrease of the hard X-ray intensity. The similarity of Cyg X-1 and the transient source A0620 00 with respect to such spectral variation has been pointed out by Coe *et al.*[1].

Cygnus X-1 was observed on November 4, 1975 from 2245 until 2330. The apparatus was launched from Palestine, Texas and floated for 5.5 h at an altitude corresponding to a pressure of 4.3 g cm⁻². The data were obtained with an array of six 100-cm² NaI (Tl) detectors which are part of an X-ray burst experiment under construction. The counts of each detector were sampled in four energy channels for equal periods of time. The data handling as well as the film recording were designed
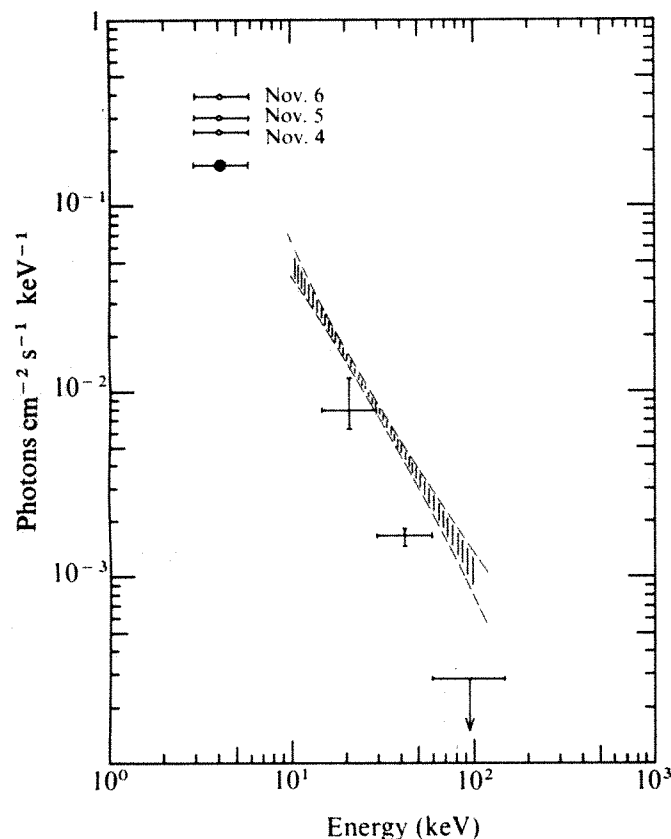
Fig. 1 Spectra of the X-ray source Cyg X-1. The black dot and hatched area indicate average spectra of Cyg X-1 in the 'normal state'. ○, Daily averages of the intensity in the 3–6 keV band for three successive days during the upward transition of Cyg X-1 (from ref. 4). The crosses are our data.

to suit the apparatus on future transatlantic balloon flights. In this short test flight we used a $4 \times 22°$ passive collimator to separate X-ray sources. The gondola was orientated by an azimuth control system. The background flux was observed shortly before and after the Cyg X-1 exposure. In the 30–60-keV range, 7×10^{-3} counts $cm^{-3} s^{-1} keV^{-1}$ were found. After subtracting the background we fitted a power spectrum to the data taking into account absorption, collimator response, energy resolution and the probability of escape of an iodine K photon. All detectors were calibrated by measuring the 60-keV line of a ^{241}Am radioactive source and the iodine escape peak at ~ 30 keV.

The resulting X-ray intensity of Cyg X-1 is shown in Fig. 1. For comparison we plotted typical normal state spectra of Cyg X-1, which are remarkably constant as often observed. The soft X-ray intensity data (from ref. 4) are daily averages for November 4, 6 and 5, 1975. Our data combined with the soft X-ray measurements suggest a single power law spectrum from 3 to 80 keV with an increasing spectral index during the upward transition of Cyg X-1 to the high state.

A power spectrum is expected if one assumes the inverse Compton effect as the basic mechanism which produces the hard X-ray tail of Cyg X-1 (ref. 2). These models require a very luminous source of soft photons inside an optically thin, hot ($kT_c \simeq 100$ keV) disk. Spectral time variation may be caused by a varying intensity of the inner soft photon source and a stable hot cloud. Spherical accretion on to a black hole has been suggested (ref. 3 and U. Anzer, G. Börner, and P. Mészáros, unpublished). Temperatures of $> 10^9$ K may be produced by the dissipative heating of the infalling gas. The X-ray spectra could then arise from bremsstrahlung which would also explain the spectra.

1 Coe, M. J., Engel, A. R., and Quenby, J. J., *Nature*, 259, 544 (1976).
2 Eardly, D. M., Lightman, A. P., and Shapiro, S. L., *Astrophys. J. Lett.*, 199, L153 (1975).
3 Mészáros, P., *Nature*, 258, 583 (1975).
4 Holt, S. S., Kaluzienski, L. J., Boldt, E. A., and Serlemitsos, P. J., GSFC-preprint-X-661-76-37 (1976).

Ball lightning as electromagnetic energy

NONE of the theories[1-8] proposed to explain ball lightning seems able to account satisfactorily for the experimental evidence[9-14]. Jennison[15] has suggested that ball lightning is a spherical stable standing wave of electromagnetic radiation. The proposal that ball lightning in effect consists of electromagnetic field energy trapped in an evacuated cavity of approximately spherical shape, with an ionised sheath separating it from the atmosphere, merits further examination.

The energy density inside the ball will be magnetic ($B^2/2\mu_0$), electric ($\varepsilon_0 E^2/2$) or both. For equilibrium, the ball must exert an average pressure at all points on its surface equal to atmospheric pressure. The standard boundary conditions for a microwave cavity with conducting walls are **E** normal and **B** tangential to the boundary. For these conditions the outward electromagnetic pressure is $[(B^2/2\mu_0)—(\varepsilon_0 E^2/2)]$. Now the largest estimate[1,9] of the energy density in a ball ($\sim 4 \times 10^9 J m^{-3}$) is 4×10^4 times the atmospheric pressure expressed as an energy density. The average magnetic energy density at all points on the surface of the ball must therefore be very nearly equal to, but slightly greater than, the average electric energy density. This is the main difficulty, but further problems include stability, power loss and mechanism of formation of the ball.

A mechanism consistent with recently reported observations[16], for the formation of a simplified 'cylindrical' ball is depicted in Fig. 1. A cloud of positive space charge is attracted towards a negative leader stroke as it approaches the ground (Fig. 1a). There is some misalignment, however, which causes parts of the two regions of space charge to spiral around each other (Fig. 1b). The end result is an electric dipole configuration rotating in the atmosphere just above the ground (Fig. 1c). This departure from the normal cloud to ground discharge mechanism requires an appropriate asymmetry in the geometry (for example, a small building or a tree[9] in the immediate vicinity). It is therefore a comparatively rare event, as required by the observed incidence of ball lightning[16]. It is well known from data on lightning return strokes[1], that once the atmosphere has been ionised, the speed of a lightning stroke can exceed $10^8 m s^{-1}$. It is an important feature of this model that the peripheral speed of the final dipolar rotating field pattern should be $> c$, the velocity of light. Given that the space charge regions are not reduced to charges of one sign only, this is quite possible, since there is then no requirement for individual charged particle velocities to be $> c$, (phase velocities $> c$ are commonplace in wave-guides and in pulsar rotating fields[17], both of which have features in common with this model).

A sketch of a rotating electric dipolar field pattern in a long cylindrical cavity is shown in Fig. 2. The return path for the current I in the wall is by way of the displacement current in the cavity which may be written as $(\nabla B \cdot \hat{Z}/\mu_0)$. The vacuum field inside the cavity may be obtained by solving the equations

$$\nabla^2 B - (\omega^2/c^2)\partial^2 B/\partial\theta^2 = 0$$

$$(\nabla B \cdot \hat{Z}) = (\omega/c^2)[r\partial E_r/\partial\theta - \partial\partial E_\theta/\partial\theta] = 0$$
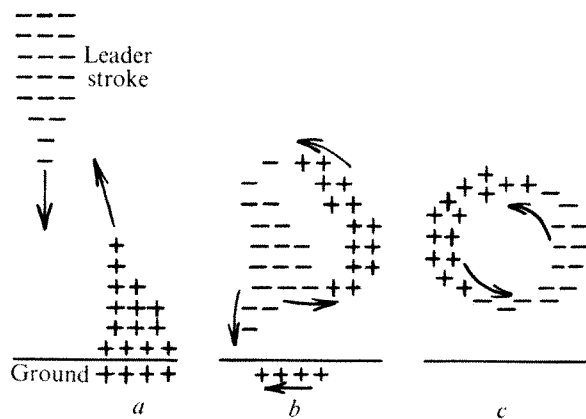
Fig. 1 Successive stages (*a–c*) in the suggested mechanism of ball lightning formation.

subject to the condition that **E** is normal to the boundary. The *r*— component of the second equation gives, for the rotating dipolar field considered

$$B = -(r\omega/c^2)E_r$$

Thus at the boundary (radius *a*), where E_θ is zero, the net outward electromagnetic pressure is

$$(B^2/2\mu_0) - (\varepsilon_0 E^2/2) = (B^2/2\mu_0)[1 - (c^2/a^2\omega^2)]$$

Therefore the peripheral velocity (*a*ω) has to be > *c* so that the net electromagnetic pressure be directed outwards. The electromagnetic pressure does of course vary with rotation angle, but the frequency is too high for the atmosphere to respond appreciably to the variations. The ratio of the electromagnetic energy density to the electromagnetic pressure at the sheath and hence to the atmospheric pressure, can be made as large as is required to fit the observations. The particles in the sheath will move at the convective velocity ($E \times B/B^2$) with a small counter-streaming velocity superimposed. In the frame moving at the convective velocity the radial electric field is zero and a small outward ($J \times B$) force balances the atmospheric pressure. Compared with the radial component of electric field in the rest frame, the tangential electric field in the sheath associated with the small counterstreaming (conduction) current will be negligible. The ball does not discharge internally because the radial electric field seen by the convected sheath is zero.

For an energy density in the ball of 4×10^8 J m^{-3}, the asso-

Fig. 2 Schematic diagram of a rotating electric dipolar field pattern in a long cylindrical cavity with conducting wall. No attempt has been made to show retardation of the field.
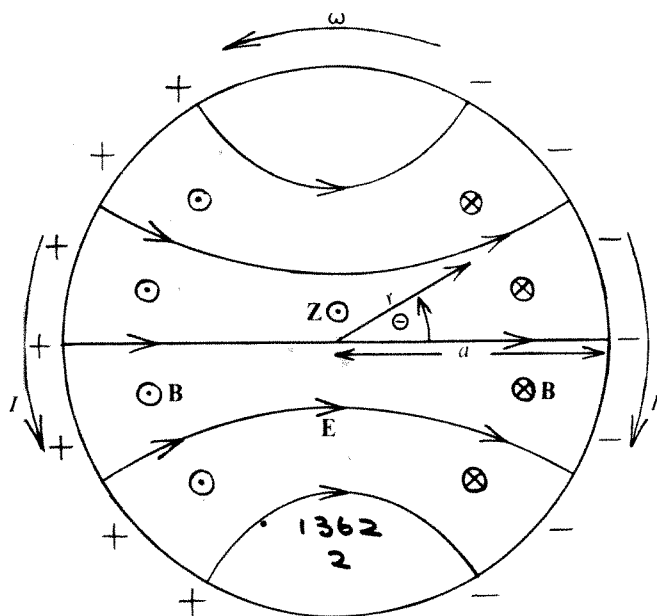
ciated values of *E* and *B* are $\sim 2 \times 10^{10}$ V m^{-1}, and 70 T respectively. For a cylinder of ~ 10 cm diameter[9] and 30 cm length, the total energy is $\sim 10^7$ J, that is, 1–10% of the energy available in a typical lightning discharge; and the charge is 0.02 C, which is $\sim 10\%$ of a typical leader charge[1], suggesting that only a fraction of the leader charge goes into formation of the ball, the remainder discharging to ground in the normal way (Fig. 1*b*).

In principle there is no difficulty in extending this model to the spherical case. To avoid nodes of zero electromagnetic pressure on the rotation axis, additional precessional velocities need to be introduced. Any air penetrating the surface of the ball near the rotation axis will be rapidly ionised and expelled. The forces required are likely to be provided by a secondary flow of conduction and convection currents across the rotation axis, so that the ball effectively generates its own precessional mechanism. The average electromagnetic pressure can then be made the same at all points on the surface, as required for equilibrium. This field computation would, however, be very complicated. Nevertheless, the most stable configuration is the most likely to occur in practice, and is likely to be approximately spherical, although any rapidly precessing volume of revolution would be indistinguishable from a sphere by the naked eye. The exact shape to be expected would depend on detailed energy considerations. The MHD instabilities that occur in plasma confinement experiments are not likely to occur with this model, because the field is inside rather than outside so that the boundary curvature is everywhere favourable, and also because there will undoubtedly be a high degree of shear in the boundary magnetic field. The wall current is likely to be largely convective, with only small conduction currents needed to keep charge velocities < *c*. Power loss through ohmic heating of the sheath is therefore expected to be small.

There is also likely to be a small leakage of the field through the sheath, which could be sufficient to maintain the ball at a steady distance from surrounding structures, but insufficient to cause motion in small metal objects nearby[15]. On the other hand, strong interaction with metal objects penetrating the surface of the ball would be expected[14]. The energy in the ball could either be entirely dissipated by ohmic heating of the sheath and radiation through it, in which case the ball would disappear quietly, or it could be dispersed by a perturbation leading to an instability, in which case the ball would disintegrate explosively. Both explosive and non-explosive terminations have been observed[4]. There is no reason, according to this model, why ball lightning should not drift into buildings[11], or be found in aeroplanes[10], since it transports its own energy. In the latter case, the ball would either have to pass through a glass window, or be formed inside the aeroplane. If in fact lightning were to strike an aeroplane near one of its windows, it is quite possible that sufficient electric field and ionisation would be set up inside the aircraft for ball formation to occur as described here. The metal surrounding the window would take the place of the ground, and the geometry would favour the necessary spiralling motion.

Undoubtedly this model will require modification as further characteristics of ball lightning become established. It does suggest, however, that the electromagnetic field inside the ball is likely to be the energy source rather than nuclear reactions[1], antimatter[7], continuous energy input from an external field[5,6], or unusual forms of gas discharge energy[2,3].

V. G. Endean

University of Durham,
Department of Engineering Science,
Science Laboratories, South Road,
Durham DH1 3LE, UK

1 Altschuler, M. D., House, L. L., and Hildner, E., *Nature*, **228**, 545–6 (1970).
2 Stakhanov, I. P., *Sov. Phys. tech. Phys.*, **19**, 861–4 (1975).
3 Neugebauer, T., *Fiz. Szle.*, **25**, 49–56 (1975).
4 Meek, J. M., and Craggs, J. D., *Electrical Breakdown of Gases*, ch. 5 (Oxford University Press, 1953).
5 Kapitsa, P. L., *Dokl. Akad. Nauk SSSR*, **101**, 245–8 (1955).

[6] Kozlov, B. N., *Sov. Phys. Dokl.*, 20, 261–2 (1975).
[7] Ashby, D. E. T. F., and Whitehead, C., *Nature*, 230, 180–2 (1971).
[8] Wooding, E. R., *Nature*, 239, 394–5 (1972).
[9] Anderson, F. J., and Freier, G. D., *J. geophys. Res.*, 77, 3928–30 (1972).
[10] Jennison, R. C., *Nature*, 224, 895 (1969).
[11] Stenhoff, M., *Nature*, 260, 596–7 (1976).
[12] Crawford, J. F., *Nature*, 239, 395 (1972).
[13] Fleming, S. J., and Aitken, M. J., *Nature*, 252, 220–1 (1974).
[14] Blair, A. J. F., *Nature*, 243, 512–3 (1973).
[15] Jennison, R. C., *Nature*, 245, 95 (1973).
[16] Rodney, P. F., and Tompkins, D. R., *Bull. Am. phys. Soc.*, 20, 659 (1975).
[17] Endean, V. G., and Allen, J. E., *Nature*, 228, 348–50 (1970).

Radiocarbon chronology of late Quaternary lakes in the Arabian Desert

A GOOD deal of information on late Quaternary moist periods in East Africa and Australia, derived largely from the dating and analysis of lake sediments, has been collected over recent years[1–4]. Although fossil lake beds have long been known to exist in Arabia[5–7], they have received scant attention. I report here the results of a preliminary attempt to establish the age of a series of late Quaternary lake sediments in the Rub' al Khali desert of the Arabian Peninsula. More detailed palaeoclimatic implications will be considered elsewhere later.

The Arabian Peninsula (Fig. 1) is essentially a plateau which dips gently towards the north-east and can be sub-divided into a western (Precambrian) igneous shield and an eastern province of sedimentary rocks. Alluvial sands, gravels and silts obscure a dendritic drainage system which was established in Quaternary times[7–9]. These sediments are in turn partially blanketed by wind-laid sands, especially in the eastern basins. The largest of these basins is the now extremely arid Rub' al Khali.

Lacustrine deposits are intimately associated with the aeolian and alluvial deposits and, though insignificant in dimensions, hold the key to the late Quaternary geomorphological and climatic evolution of the area. The deposits consist of calcareous and often fossiliferous marls, clays and silts. They characteristically form flat-topped benches and mounds, often on the lee of longitudinal dunes and in the 'corridors' between the dunes where recent deflation has removed the superficial deposits. Stratigraphical and morphological criteria indicate the presence of two main generations of lake beds, the older resting on alluvium, and the younger on aeolian sands which antedate the modern (red) dune sands. The lake beds seem to reflect centripetal runoff into deflated hollows.

Fossil Lake Mundafan (Fig. 1), one basin in which both sets of deposits are represented, displays 24 m of lake beds

Fig. 1 Location map of places mentioned in text.

Table 1 Radiocarbon dates from sites mentioned in text

Field no.*	Laboratory no.	Material	Age (yr b.p.)
		Series A	
MF 1	UGa-1218	Marl	17,460 ± 245
MF 2	UGa-1203	Marl	21,090 ± 420
MF 3	UGa-1217	Marl	21,280 ± 275
Site 1	I-6987	Marl	21,400 ± 450
MF 4	UGa-1202	Marl	22,345 ± 415
MF 5	UGa-1211	Marl	22,965 ± 390
MF 6	UGa-1209	Marl	23,075 ± 425
MF 7	UGa-1213	Marl	24,145 ± 400
MF 8	I-7427	Algal limestone	25,660 ± 810
Site 2	I-7447	Marl	27,160 ± 940
MF 9	UGa-1220	Marl	28,750 ± 615
MF 10	UGa-1219	Shells	29,595 ± 780
Site 3	I-6988	Shells	29,660 ± 1,400
MF 11	I-7111	Shells	36,300 ± 2,400
		Series B	
MF 12	UGa-1216	Algal encrustation	6,100 ± 70
Site 4	I-6410	Shells	6,520 ± 115
MF 13	UGa-1207	Diatomaceous marl	7,040 ± 115
Site 5	I-7307	Shells	7,160 ± 115
MF 14	UGa-1204	Marl	7,190 ± 85
MF 15	UGa-1206	Diatomaceous marl	7,265 ± 80
MF 16	UGa-1208	Shells	7,400 ± 210
MF 17	UGa-1205	Diatomaceous marl	7,770 ± 90
MF 18	UGa-1212	Marl	8,060 ± 95
MF 19	UGa-1221	Marl	8,155 ± 85
MF 20	UGa-1222	Shells	8,565 ± 110
MF 21	UGa-1214	Charcoal/ash	8,800 ± 90
		Series C	
MF 22	UGa-1215	Calcareous siltstone	11,465 ± 115
MF 23	UGa-1210	Marl	14,965 ± 195

*MF refers to Lake Mundafan samples. Sites of other samples are indicated in Fig. 1.

interstratified with alluvial and aeolian sands. The principal sources of runoff were probably the backslope of the eastward-dipping Wajid plateau and the foreslope of the eastward-dipping Al Arid (Jabal Tuwaiq) escarpment. The lake beds of the Mundafan basin are thought to extend beneath the alluvium and aeolian sands 150 km northward along the Al Arid escarpment.

Radiocarbon dating of lake beds from the Mundafan basin and from five other smaller depressions further inside the Rub' al Khali (sites 1–5 in Fig. 1) supports the contention that two main periods of high water level occurred in the area during the late Quaternary. The dates (Table 1) place the first period (series A) between ~36,000 and 17,000 yr b.p., with a concentration of dates between 21,000 and 30,000 yr b.p., and the second period (series B) between ~9,000 and 6,000 yr b.p. A third period of minimal and low-lying lake bed deposition (series C) separates the two.

At Lake Mundafan, extensive wind erosion, local slumping, possible changes in basin configuration, and the fluctuating nature of the various water levels combine to complicate stratigraphic relationships of the beds from which the dated material derives. Details remain to be worked out. Generally, beds of Series A occupy high parts of the basin and beds of series B and series C occupy low parts.

The dates reported here are in close agreement with dates reported from East Africa and Australia[1–4]. Additional dates and further analysis of Arabian lake bed material should help define more precisely the late Quaternary palaeoclimatic sequence of the area.

H. A. McClure

Department of Geography,
University College,
Gower Street,
London WC1E 6BT, UK

[1] Butzer, K. W., Isaac, G. L., Richardson, J. L., and Washbourn-Kamau, C., *Science*, 175, 1069–1076 (1972).
[2] Gasse, F., Fontes, J-C., and Rognon, P., *Palaeogeogr. Palaeoclimatol. Palaeoecol.*, 15, 109–148 (1974).
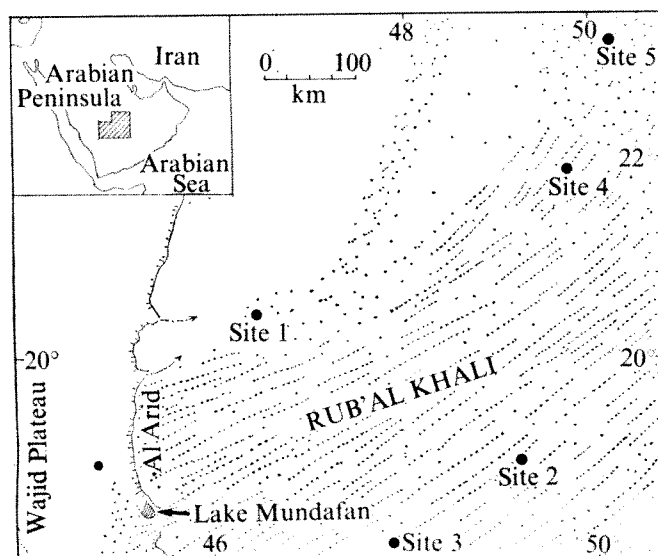
3 Grove, A. T., Street, F. A., and Goudie, A. S., *Geogrl J.*, **141**, 177–202 (1975).
4 Bowler, J. M., and Hamada, T., *Nature*, **232**, 330–332 (1971).
5 Philby, H. St. J., *The Empty Quarter* (H. Holt, New York, 1933).
6 Geology of the Western Rub' al Khali Quadrangle, *U. S. Geol. Survey Misc. Geologic Invest. Map I-218A* (1963).
7 Powers, R. W., Ramirez, L. F., Redmond, C. D., and Elberg, E. L., Jr, *U. S. Geol. Survey Prof. Pap. 560-D* (1966).
8 Holm, D. A., *Science*, **132**, 1369–1379 (1960).
9 Chapman, R. W., *Bull. geol. Soc. Am.*, **82**, 2713–2727 (1971).

Effect of magnetic field on reduction of nickel oxide

"In 1872, US patent No. 133099 on the use of a magnetic coil for 'Improvement in Reduction of Ores, etc.' was granted to Abraham T. Hay, of Burlington, Iowa[1]. It was a failure." Thus stated Skorski[2] who a century later reported that the reduction rate of haematite (Fe_2O_3) with H_2 was considerably enhanced by application of 500- and 1,400-oersted magnetic fields as compared with reduction in the Earth's field (~0.5 oersted). Skorski[2] also used CH_4 and CO for the reduction, but found the rate in these cases was slower in a strong magnetic field than in the Earth's field. He concluded that the effect must be dependent on the magnetic properties of H_2 itself rather than on those of Fe and its oxides. This notion was immediately challenged by Svare[3] who demonstrated that the magnetic properties of H_2 are not significantly different from those of CO and CH_4 in the context of the reaction involved. Peters[4], however, argued that the observed acceleration in reaction rate is predictable from thermodynamics theory when product and reactant species differ widely in their magnetic properties—as do Fe_2O_3 and metallic Fe (or NiO and metallic Ni) for example.
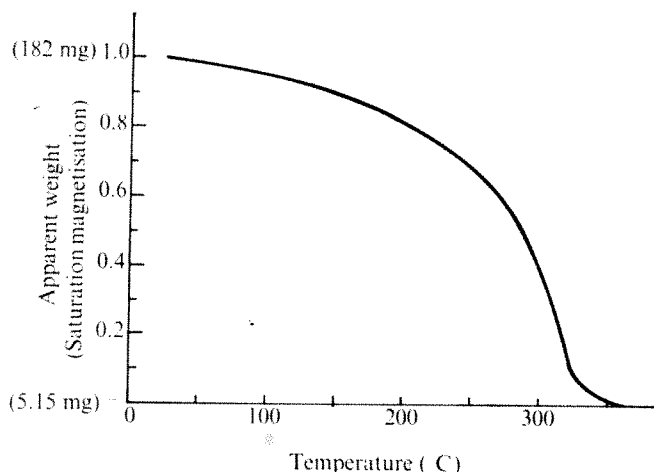
Fig. 1 Variation of the saturation magnetisation of Ni metal with temperature, as recorded by the electrobalance used in this work. Note that the apparent weight (saturation magnetisation) of Ni is much greater than its actual weight.

One can readily view this argument qualitatively in terms of Le Chatelier's principle.

We recently began a systematic study of the effect of a strong magnetic field (~4,200 oersted) on oxidation–reduction systems involving strongly magnetic products of reactants. The Fe-iron oxide system is, however, relatively complicated

$$Fe_2O_3 \longrightarrow Fe_3O_4 \longrightarrow Fe \quad (< 550\ ^\circ C)$$
haematite   magnetite   metallic iron
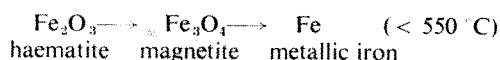
Thus Skorski[2] examined the above reduction on an 'observation after the fact' basis—the final product probably being an unknown mixture of Fe_2O_3, Fe_3O_4 and Fe. The interpretation of results is thus considerably complicated.

We present here results using a sensitive technique whereby observations are continuously recorded on a real-time basis during constant temperature reduction of NiO to Ni. Since Ni is ferromagnetic whereas NiO is paramagnetic, Peters'[4] criterion
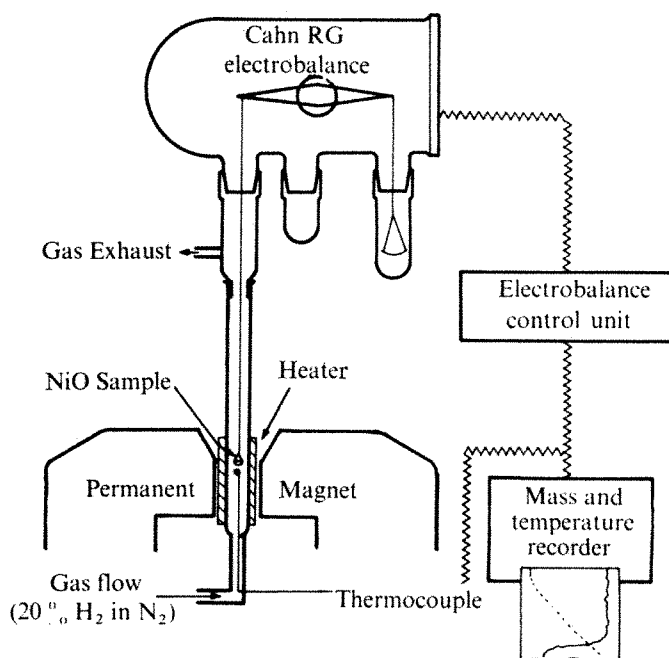
Fig. 2 Schematic diagram of the Cahn electrobalance used in this work. The system continuously records the weight of the sample as the reaction proceeds.

that product and reactant magnetic properties differ widely is met. We thus incorporated two distinctive features into our work: first, we selected a simple metal–metal oxide system; only NiO and Ni are involved. Only NiO occurs in the ordinary chemistry of nickel. Higher oxidation states of nickel are quite

Fig 3 Reduction of NiO to metallic Ni with 65 ml min^{-1} flow of 20% H_2 in N_2 carrier at 295 °C. This experiment was conducted in the Earth's magnetic field. Points were hand-read from a strip chart recording for construction of Figs 3 and 4 so they could be directly compared. The mass data were recorded continuously: □, 2.90 mg NiO; +, 3.08 mg NiO; ●, 0.92 mg NiO; and ○, 1.14 mg NiO. The solid bold line through the data points is a least squares best fit straight line fitted to the points from 0 to 20 min. It is for comparison only.
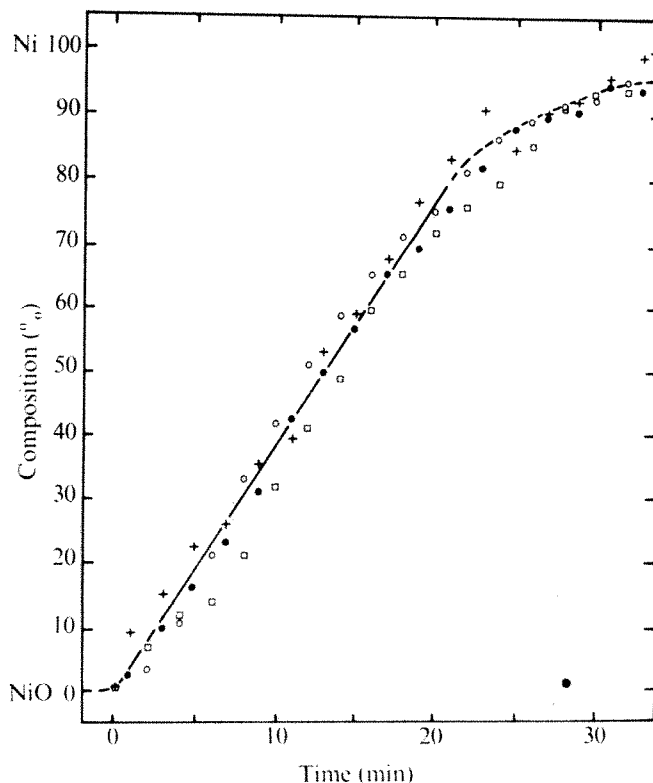
unstable; second, we continuously monitored the course of the reaction.

The behaviour of Ni in the presence of a large magnetic field ($\sim$4,200 oersted) is well known, as is the temperature dependence of this behaviour (illustrated in Fig. 1 for our system). This behaviour was determined by suspending a small sample from a sensitive balance (see Fig. 2). The application of a strong inhomogeneous magnetic field to the sample results in a force of attraction—measured as an apparent weight which is many times greater ($\sim$40 times for Ni) than the actual weight of the sample. The apparent weight (saturation magnetisation) decreases with increasing temperature until it essentially vanishes at the Curie temperature. Heating and cooling curves were indistinguishable.
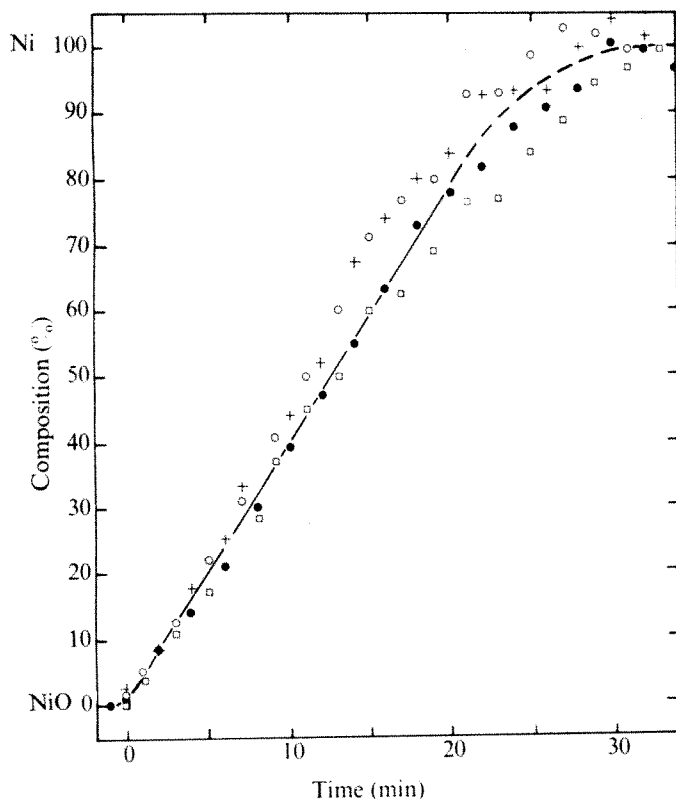
Fig. 4 Reduction of NiO to metallic Ni. Experimental conditions were the same as those in Fig. 3, except that a strong magnetic field ($\sim$4,200 oersted) was applied here. $\square$, 2.58 mg NiO; $+$, 1.29 mg NiO; $\bullet$, 2.08 mg NiO; and $\bigcirc$, 2.87 mg NiO. The line is again a least squares best fit to the data exhibited in Fig. 4 in the early portion (0–20 min).

To monitor the change in NiO and Ni during the course of the reduction of NiO without application of a magnetic field, 20% H_2 in N_2 carrier was passed through the balance system (Fig. 2 with the magnet removed) for $\sim$ 1 h to thoroughly flush the system with 1–3 mg samples of NiO in place. The temperature was then quickly raised to 295 $^\circ$C ($\sim$10 min) and the weight of the sample was continuously recorded as a function of time. The observed weight change directly reflects the relative amounts of Ni and NiO in the sample at any time. The early portion of the curve in Fig. 3 ($\sim$0–20 min) was approximately linear and the bold central line shows a least squares best fit to the average of the four runs. Each of these runs were conducted in the Earth's magnetic field ($\sim$0.5 oersted).

To ascertain the change in this reduction rate (if any) caused by the application of a strong magnetic field, we used the same experimental system (Fig. 2) but with the $\sim$4,200 oersted magnet in place to influence the sample during the course of the reduction. In this case the apparent weight change was greatly exaggerated during the course of the reduction as the non-magnetic (paramagnetic) NiO was converted into ferromagnetic Ni, which under the influence of the magnetic field

yielded an apparent weight many times the actual weight of the sample (Fig. 1). Again four independent reductions, this time carried out under a strong magnetic field, were conducted and the results are shown in Fig. 4. Again a least squares line was fitted to the approximately linear portion of the curve. The difference in slope between the line in Figs 3 and 4 was only 1.3%, well within our experimental uncertainty. It is apparent from an examination of Figs 3 and 4 that any difference in the reduction rate of NiO, with or without the application of a magnetic field, is quite small and is not discernible by our technique.

We now intend to investigate the more complex iron–iron oxide system maintaining control over reaction products by carefully mixing reaction gases so as to produce an oxygen fugacity in which the desired product is stable whereas the others are not[5].

M. W. Rowe
R. Fanick
D. Jewett
J. D. Rowe

Department of Chemistry,
Texas A & M University,
College Station, Texas 77843

1 Hay, A. T., *US Patent No. 133099* (November 19, 1872).
2 Skorski, R., *Nature phys. Sci.*, **240**, 15 (1972).
3 Svare, I., *Nature phys. Sci.*, **244**, 78 (1973).
4 Peters, C. T., *Nature phys. Sci.*, **244**, 79 (1973).
5 Larson, E. E., Watson, D. E., Herndon, J. M., and Rowe, M. W., *Earth planet. Sci. Lett.*, **21**, 345 (1974); *ibid.*, **27**, 101 (1975); *ibid.*, **29**, 283 (1976).

Charge transport by solid particles in liquid dielectrics

ONE of the more controversial aspects of the study of liquid dielectrics has been the role of natural particles in the mechanism of charge transport (see, for example, the recent review by Gallagher[1] and references therein). Theoretical estimates of the charge carried by individual particles are discrepant by several orders of magnitude. However well purified and filtered a liquid sample may be, as soon as it is subjected to electrical stress such particles may be seen in motion between the electrodes by observing scattered light from them through a microscope. They are of indeterminate composition and $\sim$ 1 μm in diameter. In the first attempt to make a direct study of charge transport by natural particles[2] reliance had to be placed on a correlation involving the human senses, which was valid for low voltages and was assumed to be capable of extrapolation to higher voltages, where speeds are too great for such methods to be used. Also, since the particles exhibit a dwell-time on the electrodes which is very much greater than the transit time the experimental problem of capturing the charge waveform is considerable. Here we report preliminary results of an attempt to correlate the waveforms formerly attributed to solid particles with physical evidence of actual particle motion obtained by observation of the light scattered from them.

As a problem in measurement this is one of intriguing complexity. It is necessary to capture two fast waveforms, a charge signal of the order of femto coulombs and a light signal of the order of a tenth of a lumen. These signals occur at random and are corrupted by noise. Indeed, the first attempt to be made to record the motion of a particle by means of scattered light failed. This involved focusing the image of the centre of the gap on to the cathode of a photomultiplier, the image of the electrode surfaces being masked off. Unfortunately, the
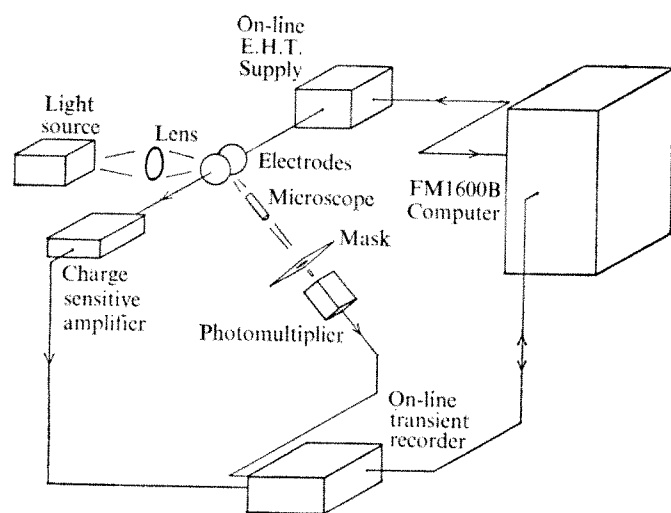
Fig. 1 Experimental arrangement for the simultaneous monitoring of motion and charge transport by natural particles. The 1-cm diameter spheres are 100 μm apart, and light scattered through 50° is collected by the microscope. The mask at the real image surface selects one electrode surface for transmission to the photocathode. The light and charge waveforms are captured by the on-line transient recorder and transmitted to the computer.

combination of high speed and low level of scattered light produced a signal which was irretrievable from noise. It was found, however, that if the image of one electrode surface was projected on to the photocathode, the arrival and departure of particles could be recorded as step waveforms in the light scattered at the electrode.

Fig. 2 Captured waveforms from the on-line incremental plotter. *a*, A typical step produced by a particle leaving the electrode. There is a slight delay from noise filtering. *b*, The corresponding charge waveform. The decay following the step rise is introduced artificially. *c*, The result of averaging 36 charge waveforms. These were adjusted horizontally to maximise registration of the corresponding light waveform steps.

Scattered light waveform

Charge waveform

Cumulatively averaged charge waveform
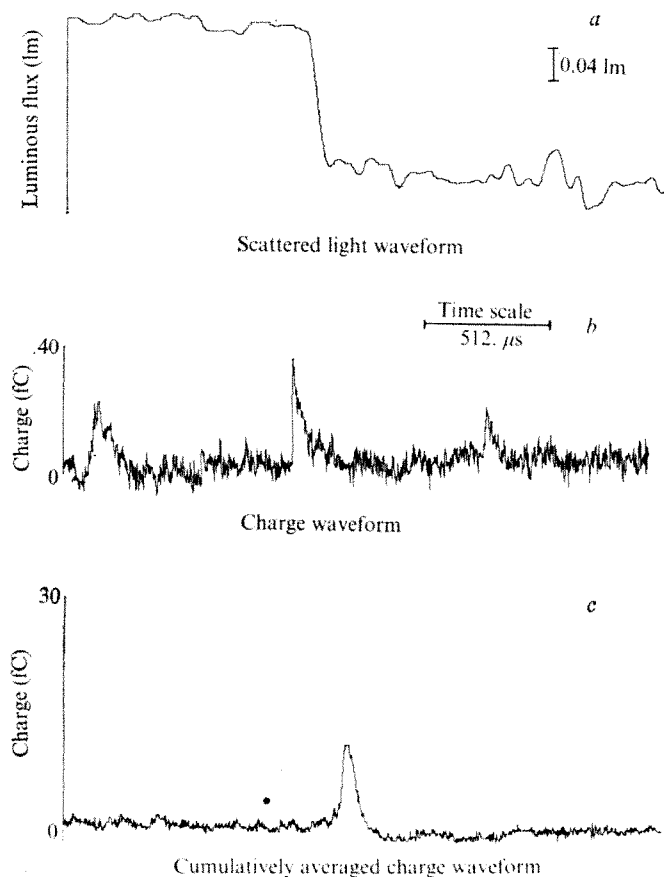
Figure 1 shows the experimental arrangement in diagrammatic form. The photomultiplier is placed at an angle of 50° to the axis of incident light and the mask is set so that scattered light from only one of the pair of spherical electrodes is recorded. Meanwhile a charge-sensitive amplifier is used to monitor the increments of charge at the cathode (though as is the practice with nuclear detectors, an artificial decay is provided in the response to prevent the 'buildup' which would saturate the amplifier). The residual problem of recording and correlating these randomly occurring waveforms was until recently a formidable one, but the introduction of the on-line transient recorder[3-5] has made it quite simple.

Figure 2*a* and *b* shows scattered light and charge waveforms recorded simultaneously by the transient recorder and plotted by the on-line computer. It will be seen that there is a coincidence suggesting that the characteristic charge transient is indeed the product of charge conveyance by a solid particle. Single instances, however, do not establish a definite correlation, and to do this the following experiment was conducted.

Pairs of waveforms like those in Fig. 2*a* and *b* were acquired repetitively at random by the on-line transient recorder. Owing to the effects of noise in the light waveforms, which are used to trigger the transient recorder, slight misalignment can occur between one light waveform and another. Therefore, within the computer the waveform pairs were shifted slightly to give best registration of the light waveforms only. Meanwhile the corresponding charge waveforms were cumulatively averaged. The sharp artificial decay is advantageous in this application since it yields a short pulse which is more difficult to register and therefore is a more demanding test for correlation. The results of 36 such repetitions are shown in Fig. 2*c*. While the shape of the charge waveform is naturally lost, its repeated correlation with the step in the scattered light establishes conclusively that particle motion is responsible for the small charge or current pulses observed in all dielectric liquids under electric stress. The speed of the equipment used allows this correlation to be observed up to stresses of 46 MV m^{-1} in *n*-hexane.

The strong grounds for accepting that the waveforms relate to particle motion are not only that other possible mechanisms are not consistent with such well defined transitions, but that the shapes are invariant from low stresses, where particles can be visually observed, through the whole stress range.

These studies confirm, first, that particles are directly responsible for much of the randomly fluctuating component of current in stressed liquid dielectrics, but second, that the charge carried by individual particles is very low and corresponds to a potential of only ∼ 15 V (refs 1, 2, 4). Perhaps even more significantly in terms of physical processes, they have not yet revealed any correlation between particle activity and the large (pC) pulses which have been observed by several authors to occur irregularly at high stresses. The ill-defined shape and duration of such pulses makes them difficult to study, even with the powerful on-line techniques mentioned above, and such a study is the object of our present work.

C. J. Buffam
J. E. Brignell

Department of Electrical and Electronic Engineering,
The City University,
London EC1V 4PB, UK

1 Gallagher, T. J., *Simple Dielectric Liquids* (Oxford University Press, 1975).
2 Rhodes, G. M., and Brignell, J. E., *Proc. 4th Int. Conf. Conduction and Breakdown in Dielectric Liquids*, 116-119 (Typografia Hiberniae, 1972).
3 Brignell, J. E., and Rhodes, G. M., *Laboratory On-Line Computing* (Intertext, London, 1975).
4 Buffam, C. J., and Brignell, J. E., *Proc. 5th Int. Conf. Conduction and Breakdown in Dielectric Liquids*, 45-48 (Delft University Press, 1975).
5 Buffam, C. J., and Brignell, J. E., in *On-Line Computers in Laboratory Use* (edit. by Rosner, R. A., Penney, B. K., Clout, P. N.), (Advanced Publications, London, 1976).

Three modes of dissociation of H bonds in hydrogen-bond dominated solids

ICE, paper and regenerated cellulose, Nylon (unstretched and lightly stretched), and to a certain extent natural cellulose, lignin and wool are typical examples of hydrogen-bond dominated solids (whose mechanical properties are mostly controlled by the density and characteristics of the hydrogen bond). Previously[1] I have shown that for these materials, Young's modulus E is related to N, the effective number of H bonds per cm³ responding to unidirectional stress, by $E = kN^{1/3}$. For cellulosics k is ~ 8×10^2 when E is in Pa. Two major rheological characteristics of hydrogen-bond dominated solids are their ready softening by water and their relaxation of stress with time under constant strain. Thus, E_w/E_0, the ratio of the modulus at w g H_2O/g solid to the modulus at $w = 0$, drops to very small values for paper at saturation. Similarly, E_t/E_0, the modulus at time t to E at $t = 0$ in a stress relaxation experiment, drops to small ratios at high values of t. I postulate that both types of loss in E arise from reductions in N, through H-bond dissociation. I find this dissociation can be in one of three modes: I, II and III.

Mode I dissociation occurs on wetting a hydrogen-bonded solid within the range, $0 < w < w_c$ where w_c is a critical moisture content designating transition from mode I to mode II. w_c is postulated to be equal to w_m, the moisture content for a monomolecular layer calculated by BET equations[17] applied to the adsorption isotherm of water on the particular polymer. Mode II occurs when $w_c < w < w_{sat}$ where w_{sat} denotes the water saturation value of w. Further, let N_0 be the value of N at $w = 0$ and W be the hypothetical value of w which allows one H_2O molecule for each potential H bond in the solid. (For cellulosics, with potentially six H bonds existing per repeat unit of cellobiose, $W = 1/3$.)

In mode I we begin with

$$d(N_w/N_0)_1 = -(N_w/N_0)_1 d(w/W) \quad (1)$$

and arrive at

$$\ln(E_w/E_0)_1 = (-1/3)(w/W) \quad (2)$$

(The suffix 1 denotes mode I.) Hence, specifically for cellulosics

$$\ln(E_w/E_0)_1 = -w \quad (3)$$

This has been demonstrated for paper for $0 < w < w_m$ (ref. 1).

In mode II, dissociation is cooperatively undertaken by a number of bonds breaking down together. I use here Frank and Wen's hypothesis[2] of H bonds making and breaking co-operatively together in flickering clusters. Thus, I postulate that a molecule of water initiates the cooperative phenomenon by breaking only one bond, but this triggers others to break immediately. The total number is equal to the cooperative index (CI) but the reaction is still unimolecular in N. Thus

$$d(N_w/N_0)_2 = -(\overline{CI})(N_w/N_0)_2 d(w/W) \quad (4)$$

with initial condition at $w = w_c$

$$\ln(N_w/N_0)_2 - \ln(N_w/N_0)_1 = -(w_c/W) \quad (5)$$

Hence

$$\ln(E_w/E_0)_2 = (1/3)\{(w_c/W)[(\overline{CI}) - 1] - (\overline{CI})(w/W)\} \quad (6)$$

For cellulose, $W = 1/3$, and equation (6) reduces to

$$\ln(E_w/E_0)_2 = w_c[(\overline{CI}) - 1] - (\overline{CI})w \quad (7)$$

The mean cooperative index $(\overline{CI})$ can be calculated from rheological data, but an *a priori* value was obtained from Starkweather's analysis[3] of the cluster behaviour of water adsorbed on polymers, using Zimm's thermodynamic concepts[4,5], as well as calculations by Némethy and Scheraga[6] for water at room temperature and found, for cellulose and Nylon 66, to be: $(\overline{CI}) = 6.71$.

Comparison of data in the literature on the dependence of E_w on w for paper[7-11] gives the mean for $(\overline{CI})$ as 6.70 ± 0.89 The mean w_c for these experiments was 0.047 ± 0.011, whereas w_m for paper is normally between 0.04 and 0.06. The average for different cellulosics studied by Meredith[12] (ramie, Fortisan, viscose rayon and mercerised cotton) was found to be $(\overline{CI}) = 7.7 \pm 2.7$. The value for w_c was indeterminate because of the large standard deviations in the data, (-0.042 ± 0.09). The average for Nylon 66 studied by Meredith[12] and others[13,14] was $(\overline{CI}) = 6.70 \pm 2.1$. The average w_c was indeterminate but clearly < 0.01 g H_2O/g solid $(0.003 \pm 0.005$; W for Nylon 66 is 0.1593).

Thus, the cooperative dissociation of H bonds in mode II is still 'unimolecular' in N, even though approximately seven bonds are triggered to break down together on average.

In mode III, controlling stress relaxation under constant strain, the dissociation is polymolecular in N. This may be explained by imagining the need for a minimum number of bonds, α, having to dissociate simultaneously before a unit jump leading to stress relaxation becomes possible. Thus, it was found empirically from analysis of rheological data[1]

$$\frac{d(N_t/N_0)}{dt} = -\sum_{i=1}^{i=\prime} k_i (N_t/N_0)^{\alpha_i} \quad (8)$$

(where $N_0 = N$ at $t = 0$) with explicit interconnections between k_i with k_{i+1} and α_i with α_{i+1} respectively. The infinite series of different reaction rate constants k_i and cooperative indices α_i soon degenerate into a single reaction of rate k_1 and of order α_1, the 'principal cooperative index', which controls the stress relaxation curve for all values of $(N_t/N_0) \lesssim 0.8$.

Analyses of data published by Houghton and Sellen[15] on dry and wet regenerated cellulose at subzero temperatures reveal interesting maxima and minima in α_1 at particular temperatures.

Thus, in mode III the cooperative dissociation of H bonds is different from that in mode II in being truly polymolecular whereas in mode II it is unimolecular in N.

Details of these studies will be published elsewhere[16].

ALFRED H. NISSAN

Westvaco,
299 Park Avenue,
New York, New York 10017

1. Nissan, A. H., *Trans. Faraday Soc.*, 53, 700–721 (1957); *Nature*, 175, 424–425 (1955); in *Surfaces and Coatings* (edit. by Marchessault, R. H., and Skaar, C.), 221–265 (Syracuse University Press, 1967).
2. Frank, H. S., and Wen, W-Y., *Discuss. Faraday Soc.*, 24, 133–140 (1957).
3. Starkweather, H. W., Jr, *Macromolecules*, 8, 476–479 (1975).
4. Zimm, B. H., *J. chem. Phys.*, 21, 934–935 (1955).
5. Zimm, B. H., and Lundberg, J. L., *J. phys. Chem.*, 60, 425–428 (1956).
6. Némethy, G., and Scheraga, H. A., *J. chem. Phys.*, 36, 3382–3400; 3401–3417 (1962).
7. Higgins, H. G., and Balodis, Y., *Fracture* (edit. by Osborn, C. J.), 201–220 (University of Melbourne Press, 1965).
8. Andersson, O., and Berkyto, E., *Svensk Papp Tidn.*, 54, 441–444 (1951).
9. Benson, R. E., *Tappi*, 54, 699–703 (1971).
10. Riemen, W. P., and Kurath, S. K., *Tappi*, 47, 629–633 (1974).
11. Brezinski, J. P., thesis, Institute of Paper Chemistry, Appleton, Wisconsin (1955).
12. Meredith, R., *J. Text. Inst.*, 48, T163–T174 (1957).
13. Adams, N., *J. Text. Inst.*, 47, T530–T540 (1956).
14. Quistwater, J. M. R., and Dunell, B. A., *J. polym. Sci.*, 28, 309–318 (1958); *J. appl. polym. Sci.*, 1, 267–271 (1959).
15. Haughton, P. M., and Sellen, D. B., *J. Phys. D.*, 6, 1998–2011 (1973).
16. Nissan, A H., *Macromolecules* (in the press).
17. Brunauer, S., Emmett, P. H., and Teller, E., *J. Am. chem. Soc.*, 60, 309 (1938).

Cumulative deformation of magnesium oxide crystals by softer sliders

MECHANISMS postulated for the wear of hard crystalline surfaces by softer solids include thermal degradation[1], diffusion and solution[2]; and self abrasion of the harder solid when its wear debris is picked up by the softer surface[3]. A mechanism based purely on mechanical aspects has not yet been established but there has been the suggestion of an 'incubation period' during which the structure of the hard solid has been modified before the onset of visible wear[4]. In earlier publications[5,6], we have established that a dislocated zone was formed in some hard crystals by a softer slider, whenever the pressure between the contacting surfaces was greater than a certain threshold value which was thought to be related to the critical resolved shear stress. For a given load, and in these conditions, the depth of this zone was independent of the slider hardness—regardless of whether the point itself was blunted or visible macroscopic deformation and penetration of the harder crystal was obtained. Then, the depth of the dislocated zone was determined only by the magnitude of the applied load. There were two principal objectives to this particular study. First, to measure the effect of slider hardness on the number of traversals necessary to initiate visible fragmentation and wear in the harder crystal. Second, to use dislocation etching and microhardness techniques to follow the deformation of the harder crystal during the microdeformation preceding fracture—the incubation period.

a

b

c
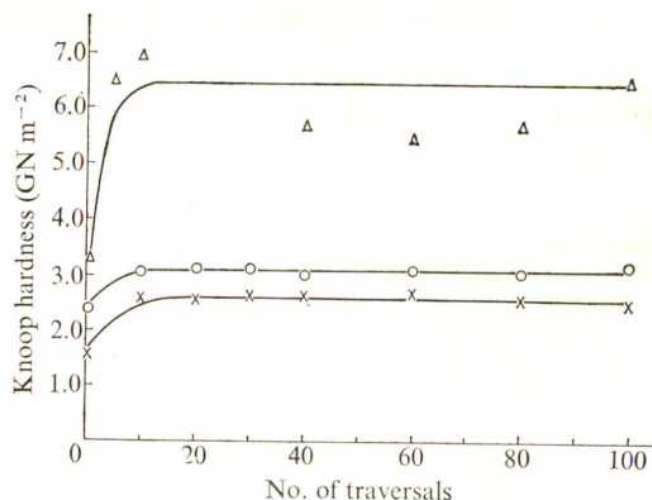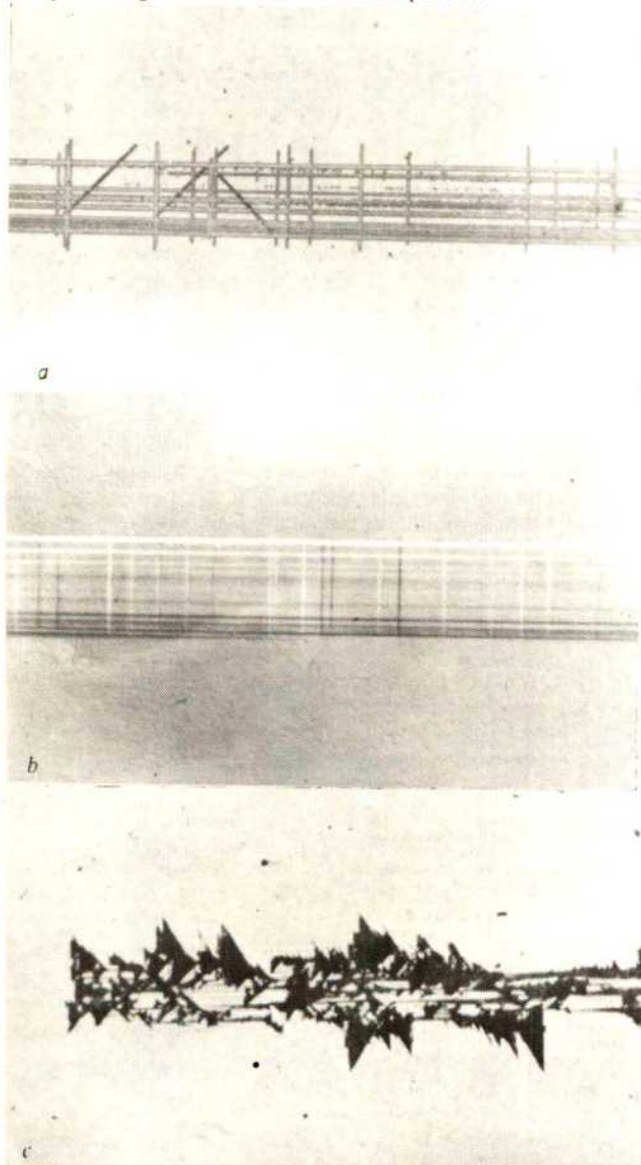
Fig. 1 Illustrating the degree of work hardening, as the tip of the slider was blunted, during the initial traversals on MgO, for △, EN 58J; ○, 95/5 phosphor bronze; ×, 70/30 brass.

The apparatus used has been described in detail elsewhere and only brief experimental details will be given here. All the following results are based on experiments using magnesium oxide as the hard crystal with conical sliders, having included angles of $136°$ at the start of the experiment, made from polycrystalline alloys. The alloys selected for this purpose included an austenitic stainless steel (EN 58J), an α brass (70/30), and a phosphor bronze (95/5). Experiments were carried out at room temperature, using reciprocating sliding at a speed of 3×10^{-4} m s^{-1}, and the surfaces were continuously lubricated with soluble or mineral oils. The direction of sliding was aligned with $\langle 100 \rangle$ crystallographic directions on the chemically polished (001) plane of MgO and the length of the resultant wear track was ~ 0.01 m. Normal loads of 0.5, 1 or 2 kg were used.

The first few traversals caused a flattening of the tip on the metallic slider with a corresponding increase in hardness from conventional work-hardening processes. This effect is reflected in microhardness measurements on the blunted tip, using a Knoop indentor with a normal load of 100 g, as a function of the number of traversals (Fig. 1). The hardness of the slider tended to a constant value, presumably indicating the ultimate level of work hardening for that particular material, within the first ten cycles. There was no visible damage on the MgO surface after the first traversal but dislocations in the contact zone, induced by the sliding, were revealed by etching (Fig. 2a). Subsequent traversals produced slip lines on the MgO surface—typically within the first ten cycles (Fig. 2b). Eventually, a visible crack was formed and then was immediately followed by significant fragmentation (Fig. 2c). The crack pattern was of the chevron form, characteristic of rocksalt crystals, produced by fracture on {110} and {100} planes[7]. The number of traversals to cause the initiation of the first crack, and thus fragmentation, was related to the ultimate hardness of the slider material. The harder the slider, the fewer the number of traversals (N_c) required to produce a crack.

Fig. 2 Stages in the deformation produced by a lubricated slider, under a normal load of 500 g, making traversals in a [100] direction on the (001) plane of MgO. a, Etch pits revealing dislocations after first traversal, ($\times 160$). b, Slip lines visible after ten traversals. ($\times 160$). c, Fragmentation and wear marking the end of the experiment at $N_c = 15,000$. ($\times 24$).

A plot of the ultimate hardness of the slider (H_s) against the critical number of traversals to fracture (N_c) gives a curve which resembles that of an S:N curve based on conventional fatigue data (Fig. 3) in which the maximum stress applied by the slider in each traversal is directly related to the hardness of the slider material. It is important to note, in this context, that the critical number of cycles was independent of the normal load for a given slider. Also, that directly comparable results were obtained with a mineral oil and a water-soluble lubricant. Thus it would appear that this particular change in environment is not important in the wear mechanism. Further support for a mechanism based on fatigue criteria was given by an experiment in which an MgO surface was abraded with a brass for ~ 12,000 cycles—that is, well after maximum work hardening of slider and crystal but before the onset of fragmentation and wear. Subsequently, the top 5 μm of the dislocated zone, which penetrated the crystal to a total depth of 150 μm, was removed by chemical polishing. Sliding was then continued, along the same track, and a further 15,000 traversals were required to produce fragmentation and wear. This compared with $N_c = 15,000$ for the original, undeformed, surface. Note that each point on the curve in Fig. 3 represents the average of at least six measurements, all within a range of 10%.

The coefficient of friction (μ) was measured throughout the experiments for all three types of slider. Where the lubricant became ineffective, adhesion caused a marked increase in friction and fragmentation was immediately observed. Results from such experiments have not been included in Fig. 3. In the normal case, μ decreased during the initial mutual work hardening of slider and crystal but then remained constant until the onset of fragmentation. At this point, an immediate and significant increase in friction was observed. It can, however, be concluded that fragmentation was responsible for the increased friction and not vice versa.

Examination of the abraded MgO surface, using dislocation etch pit techniques and microhardness measurements, confirmed some of the earlier observations made on the basis of a single
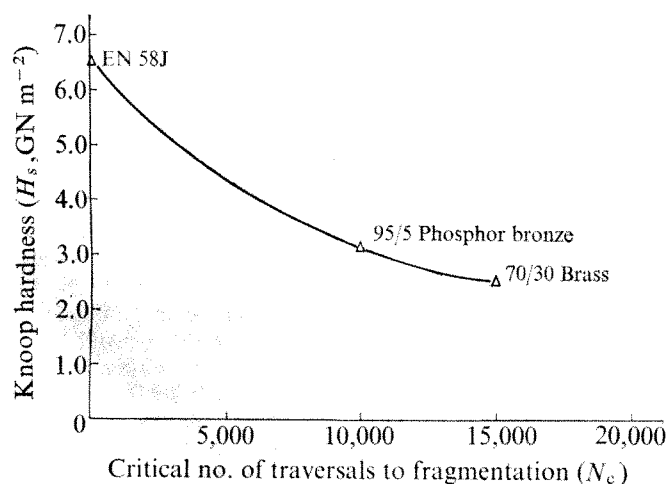
Fig. 3 The H_s:N_c curve for a (001) MgO surface abraded in the [100] direction. Note a limiting slider hardness of ~ 2.25 GN m^{-2}.

indentation or traversal[5] and gave further information on the nature of the incubation period. The depth of the resultant dislocated volume was determined not by the shape or hardness of the slider but only by the magnitude of the applied load. Dislocations produced by repeated traversals were generally contained within the dislocated volume established by the first traversal. Thus the dislocated volume remained constant whilst the dislocation density increased during the first hundred traversals. Furthermore, microhardness measurements show that the increase in dislocation density causes work hardening both on the surface and within the bulk of the magnesium oxide. Typical values, using a Knoop indenter aligned in a [110] direction on the (001) MgO surface with a load of 100 g, are shown in Fig. 4 for all three slider materials. It would seem that
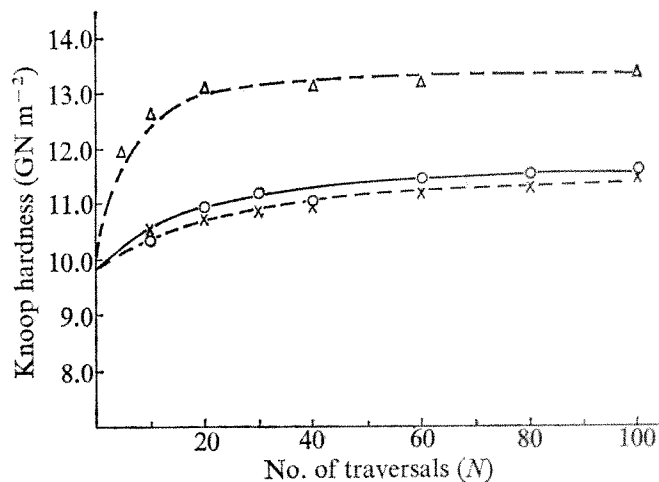
Fig. 4 Hardening of the MgO surface, by traversals on the (001) plane in a [100] direction, was shown to be dependent on the ultimate hardness of the slider material. △, EN 58J; ○, 95/5 phosphor bronze; ×, 70/30 brass.

the degree of work hardening in the MgO is again related to the ultimate hardness of the slider material. Results obtained by removing the surface layers, using chemical polishing techniques, and indenting on the newly exposed surfaces in the same manner reveal the hardening in depth below the original surface (Fig. 5).
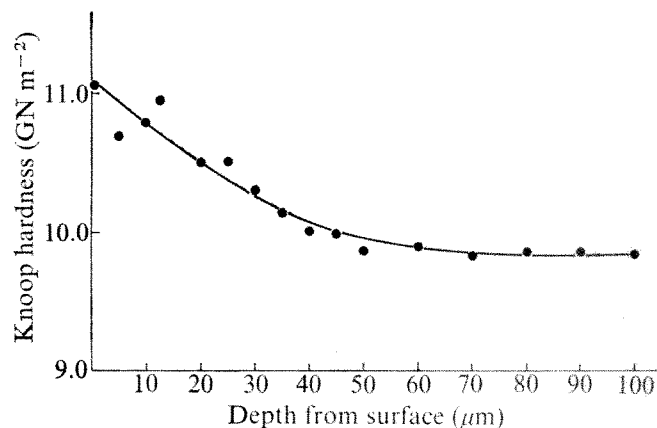
Fig. 5 Bulk hardening of the MgO crystal after ~ 200 traversals of brass slider under a load of 500 g.

In summary, these experiments provide a method of identifying the range of slider hardness within which fragmentation and wear of a harder crystal can be anticipated. Also, we now have the possibility of establishing a limiting slider hardness, akin to a fatigue or endurance limit, below which fragmentation of a given solid will not take place in a given lifetime. For example, MgO with a hardness of 9.8 GN m^{-2} should have a reasonable lifetime, in abrasive conditions, provided the hardness of the softer material does not exceed 2.25 GN m^{-2}. Current work has established that the behaviour of various other non-metallic crystals is comparable with MgO and similar H_s:N_c curves have been obtained. Clearly this information is of technological interest wherever a soft material abrades a harder one—as in gemmological applications and metal cutting operations.

C. A. Brookes
M. P. Shaw

*Department of Engineering Science,
University of Exeter, Exeter, UK*

1 Bowden, F. P., and Scott, H. G., *Proc.R.Soc.*, A258, 268 (1958).
2 Trent, E. M., *Proc.R.Soc.*, A212, 467 (1952).
3 Bowden, F. P., and Tabor, D., *Friction and Lubrication of Solids*, Part 2, (Oxford Clarendon Press, 1964).

4 Crompton, D., Hirst, W., and Howse, M. G. W., *Proc.R.Soc.*, A333, 435 (1973).
5 Brookes, C. A., and Green, P., *Nature phys.Sci.*, 246, 119–122 (1973).
6 Brookes, C. A., Green, P., and Harrison, P. H., *Diam. Res.*, 11, (1974).
7 Bowden, F. P., and Brookes, C. A., *Proc.R.Soc.*, A295, 244 (1966).

Feathering and flight evolution in *Archaeopteryx*

WITH the description of further specimens[1,2] there has been a resurgence of interest in *Archaeopteryx*. Although there is no final agreement on whether it was mainly arboreal or terrestrial, recent opinion[3] seems to favour the latter. There has been discussion on its reptile origins[4–6] and on its ability to fly and the general problem of the evolution of powered flight[3,7–12]. However, most of these studies have been made by specialists who are not ornithologists and who usually tend to extrapolate from reptile forms and behaviour in their discussions. Whatever its precursors, *Archaeopteryx* is regarded as a bird and to interpret adequately such evidence as we have we must consider it in relation to what we know of birds as well as reptiles.

Archaeopteryx possessed structures very closely resembling the flight feathers of Recent birds, which should confer an advantage by supporting it during movement, while body feathers, if present, would aid thermoregulation; but feathers also have disadvantages in being fragile and subject to abrasion. In Recent birds this is partly overcome by regular moult and replacement, but we do not know if *Archaeopteryx* had a moult. Recent birds also avoid feather damage by postures and movements in which the plumage does not usually rub or strike against surrounding objects, the larger feathers of wings and tail usually being held clear of the substrate. The possession of feathers by *Archaeopteryx*, even were it a reptile, should therefore impose some limitations on both posture and movement, comparable with those of Recent birds; and these limitations would be likely to have modified some of its characters already, and to affect its subsequent evolution.

Feathers probably evolved before avian flight and Ostrom[3] has suggested that the large feathers of the forelimbs were originally produced for a different purpose. He regards the digits as functional in catching small live prey and suggests that the feather may have formed a large area of catchment for trapping small creatures such as insects. It is possible to envisage an earlier and shorter form of feathers being spread like small nets and slapped down over insects to aid capture, but I would suggest that the feathers on *Archaeopteryx* had developed beyond this stage and that to use them for such a purpose would be likely to damage and destroy them, while the total size of the spread wing would tend to hinder the use of the forelimbs for rapid movements in catching prey.

It cannot be argued that the similar large feathers along either side of the tail were used for capturing prey, and in view of their position it is reasonable to suppose that their function was for support during movement. It is also likely that the extensive feathering of the forelimbs had a support function at this stage in evolution, and while Ostrom[6,12] argues that the skeletal structure precludes an ability to fly in the true sense, there seems to be no objection to the idea that the feathers of forelimbs and tail might be used to prolong the passage through the air of an individual that had leaped from some eminence, whether by gliding or by some crude flapping flight.

The feathering should affect both normal movement and the subsequent evolution of flight. Some earlier reconstructions of *Archaeopteryx* and theoretical proavian forms based on Recent reptiles[13,14] assumed that *Archaeopteryx* climbed trees by using the well developed digits, and launched itself into the air from a height; with flight evolving through a stage of gliding down from high perches. The feathering of *Archaeopteryx* is developed to a degree that makes it most improbable that it would climb in a lizard-like fashion or clamber among twigs using all four limbs as some reconstructions suggest. Any arboreal movement is likely to have been limited to perching, running along branches and leaping between them, meanwhile probably using the wings for balance and aerial support.

In his recent papers Ostrom[3,6,12] describes *Archaeopteryx* as a cursorial, and presumably terrestrial predator. This would accord with the above comments but would not explain why powered flight should have evolved in such circumstances. He believes the flight of *Archaeopteryx* to have been "flapping leaps" to capture insect prey, and presumably envisages flight evolving as an aid to predation. In Recent birds, flight as a means of catching prey appears to be a secondary adaptation, for species which use it, from simple flycatchers to the more elaborate aerial feeders, have additional structural modifications for this purpose, and in its earlier stages flight would have had the greatest value as a means of escape from predators. Driver and Humphries[15] in their studies of protean behaviour have shown that an element of unpredictability in the behaviour of a prey species will help it to evade capture. If prey which has previously moved along the ground suddenly launches itself into the air in a long leap, leap and glide, or full flight, a terrestrial predator is temporarily baffled and likely to discontinue the hunt. This is a very effective anti-predator device, a fact reflected in the evolution of such behaviour in a number of different groups of animals. Anyone who has tried to catch an almost-fledged nestling or a slightly injured adult bird without the powers of sustained flight will be aware of the way in which even brief bursts of short flight by the prey can make capture difficult for a relatively sophisticated predator such as man.

We can envisage *Archaeopteryx* as a bird-like creature probably resembling the coucals, *Centropus* spp., which are similar in general proportions with rather weak rounded wings, long tails and well developed legs. These latter birds are mainly terrestrial or semi-arboreal, walking or running through low vegetation or travelling through shrubs and low trees, moving along branches and jumping from branch to branch. They fly with reluctance but if hard-pressed will ascend to a higher perch and then take off in a weak and frantic flight, travelling a little distance before dropping into low vegetation and disappearing.

Archaeopteryx could have moved in similar fashion, and if launching itself from some eminence enabled it to escape terrestrial predators, there would have been a selective advantage in a readiness to take flight. If the ability to move through the air was poor and likely to result in a landing too close to the point of take-off there could be a likelihood of rediscovery during random searching by the predator. There would therefore be additional selection pressure in favour of individuals that could travel further through the air before landing. It can be argued that factors such as these could provide pressures for the evolution of powered flight in a mainly terrestrial animal without the need for an ability to climb in the strict sense or for the modification of specialised predatory leaps.

Regarding *Archaeopteryx* as a bird is therefore useful in predicting the likely limitations implied by the possession of feathers, and the probable advantages of even brief periods of movement in the air. It may prove helpful in interpretating other aspects of this important evolutionary link.

C. J. O. HARRISON

Sub-department of Ornithology,
British Museum (Natural History),
Tring, Hertfordshire HP23 6AP, UK

1 Ostrom, J. H., *Proc. K. ned. Akad. Wet.*, B75, 289–305 (1972).
2 Wellnhofer, P. *Palaeontographica, Abt. A.*, 147, 169–216 (1974).

3 Ostrom, J. H., Q. Rev. Biol., 49, 27–47 (1974).
4 Walker, A. D., Nature, 237, 257–263 (1972).
5 Ostrom, J. H., Nature, 242, 136 (1973).
4 Ostrom, J. H., J. Linn. Spc. Biol., 8, 91–182 (1976).
7 Yalden, D. W., Ibis, 113, 349–356 (1970).
8 Bramwell, C. D., Nature, 231, 128 (1971).
9 Heptonstall, W. B., Nature, 228, 185–186 (1970).
10 Heptonstall, W. B., Nature, 231, 128 (1971).
11 Heptonstall, W. B., Fauna, 4, 9–16 (1972).
12 Ostrom, J. H., Smithson. Contr. Paleob., 27, 1–21 (1976).
13 Heilmann, G., The Origin of Birds (Appleton, New York, 1972).
14 Stephen, B., Urvogel (A. Ziemsen Verlag: Wittenberg, Lutherstadt, 1974).
15 Driver, P. M., and Humphries, D. A., Nature, 226, 968–969 (1970).

Intake and digestion of hill-land vegetation by the red deer and the sheep

DOMESTICATED sheep and feral red deer (*Cervus elaphus*) are the two principal producers of food for human consumption from a large proportion of the rough grazings of Scotland. There is considerable interest in the possibility of increasing the production of venison by farming the red deer, and studies of the degree of domestication required, the development of biologically sound management practices and their economic viability are in progress[1]. A fundamental question is whether the red deer can utilise the indigenous vegetation found on the hills more efficiently than does the sheep, and, if so, to what extent this is brought about by its greater ability to range over and graze selectively on hill vegetation or by its ability to consume and digest more of the poor quality vegetation. We have investigated this latter aspect and now report that on such diets red deer consume greater quantities of digestible dry matter per unit metabolic body weight than sheep. The greater ability of the deer to increase voluntary food intake during the summer months is also discussed.

For this study we used 12 castrated male red deer and 12 castrated male Scottish Blackface sheep, aged 18 months. Both species were treated in a similar manner from 6 months of age and were accustomed to being housed in metabolism pens before being given diets of common heather (*Calluna vulgaris*, L. Hull) and *Agrostis/Festuca* grass heath (two of the common types of hill vegetation grazed by both species). In two experiments carried out in January and April 1976, the animals were offered, *ad libitum*, freeze-stored vegetation which had been harvested mechanically in the autumn and which was known to provide a diet similar to that selected by grazing during the winter (unpublished data of J.A.M.). Measurements were made of the voluntary food intake (VFI) and digestibility of each diet. The particulate-phase marker, [103]Ru-phenanthroline[2], was used to estimate the mean retention time[3] (MRT) of digesta between the rumen and the faeces—an index of the time undigested food particles spend in the gastrointestinal tract. Results are given in Table 1.

In both January and April the red deer ate at least twice as much heather and *Agrostis/Festuca* as the sheep. Surprisingly, on the heather diet the deer were able to digest their higher intake to the same extent as the sheep (Table 1), but with the *Agrostis/Festuca* diet they showed a more predictable fall in digestibility (7–8 percentage units). On both diets MRT was considerably shorter for the deer than the sheep. It is not possible to assess to what extent these differences were due to the different intakes of the two species, but further experiments have shown that when the same animals consumed similar quantities of a higher quality ration (dried grass pellets), the MRT in the red deer was still significantly ($P < 0.001$) shorter than in the sheep (unpublished). This latter result is in agreement with earlier studies using chopped dried grass and grass hay diets[4,5] which suggest a real between-species difference in the speed at which digesta passes through the digestive tract.

The VFIs of both the red deer and the sheep were higher in April than in January. In the sheep, the 40% increase in VFI of heather in April was accompanied by a reduction in the digestibility of the diet and in MRT. This result is in agreement with relationships between VFI, digestibility and MRT previously found in sheep and cattle and which support the generally accepted hypothesis that with medium-quality long roughage diets, the control of VFI in sheep is related mainly to the filling effect in the digestive tract, particularly in the rumen, and to the rate of removal of food residues from that organ[6].

In contrast, the large increase in VFI of deer of 70% between January and April, which has been associated with changes in daylength[7], was not accompanied by any change in digestibility or MRT. This suggests that the amount of feed residues transferred per unit time down the gastro-intestinal tract of the red deer was higher in April than in January. The hypothesis that VFI of long roughages is controlled by the filling effect in the digestive tract can only hold for the deer at both times of the year if the volume or capacity of their digestive tract, or at least some section of it, had increased considerably. Although the mechanisms whereby this may occur are not known, it is likely that they are under some form of hormonal control related to daylength.

In spite of these seasonal differences in VFI the red deer ate twice as much digestible dry matter as the sheep on both kinds of hill vegetation. This suggests that red deer are better adapted than sheep to grazing hill vegetation in winter and spring, providing that both species exhibit similar grazing behaviour, have similar maintenance requirements per unit of metabolic liveweight, and have similar rates of production and utilisation of the end products of digestion. These latter aspects are being investigated further.

J. A. MILNE
J. C. MACRAE
ANGELA M. SPENCE
S. WILSON

Hill Farming Research Organisation,
Bush Estate,
Penicuik,
Midlothian EH26 0PY, UK

1 Blaxter, K. L., Kay, R. N. B., Sharman, G. A. M., Cunningham, J. M. M., and Hamilton, W. J., Farming the Red Deer (HMSO, London, 1974).
2 Tan, T. N., Weston, R. H., and Hogan, J. P., Int. J. Appl. Radiat. Isotopes, 22, 301–308 (1971).
3 Faichney, G. F., in Digestion and Metabolism in the Ruminant (edit. by McDonald, I. W., and Warner, A. C. I.), 277–291 (University of New England, Armidale, 1975).

Table 1 Intake and digestibility of diets

		DMI	DMI/ kgW[0.75]	DMD	MRT
		(g)	(g)	(%)	(h)
January					
Red deer	*Agrostis/Festuca*	1,199	49.1	40.1	27.4
	Heather	847	35.0	45.3	37.7
Sheep	*Agrostis/Festuca*	432	24.1	48.8	56.1
	Heather*	327	17.7	45.7	67.3
April					
Red deer	*Agrostis/Festuca*	2,028	76.2	41.9	32.5
	Heather†	1,446	56.6	45.2	42.4
Sheep	*Agrostis/Festuca*†	485	26.6	48.1	55.9
	Heather	458	24.5	42.8	56.2
s.e.m.		±74.9	±2.87	±1.06	±3.33

*Mean of four observations.
†Mean of five observations.
Figures show mean daily voluntary intake of dry matter (DMI), dry matter intake on metabolic liveweight basis (DMI per kgW[0.75]), dry matter digestibility (DMD %) and mean retention time of a particulate-phase marker (MRT) for red deer and sheep offered diets of *Agrostis/Festuca* and common heather in January and April (means of six observations)

[4] Maloiy, G. M. O., and Kay, R. N. B., *Q. Jl exp. Physiol.*, 56, 257–266 (1971).
[5] Kay, R. N. B., and Goodall, E. D., *Proc. nutr. Soc.* (in the press).
[6] Campling, R. C., in *Physiology of Digestion amd Metabolism in the Ruminant* (edit. by Phillipson, A. T.), 226–234 (Oriel, Newcastle-upon-Tyne, 1970).
[7] Simpson, A. M., thesis, Univ. Aberdeen (1976).

Superposition images are formed by reflection in the eyes of some oceanic decapod crustacea

LIKE moths, the eyes of many pelagic and bottom-living decapods appear to glow brightly when viewed from the direction of illumination[1]. It has generally been assumed that the optical system involved is similar to that in nocturnal insects, that is, the superposition mechanism of Exner[2]. The essential feature of Exner's theory is that the optical apparatus of the cornea and crystalline cone refracts light through twice the angle at which it was incident on the cornea (Fig. 1a). Light from many facets then combines to form a single image on the receptor layer, and if this is backed by a tapetum the light is reflected out through the same facets that it entered, causing the glow. There is now no doubt that Exner's mechanism is present in some insects[3]. Recent accounts have, however, cast doubt on whether it occurs in crustacea[4,5]. This is largely because the high refractive indices and inhomogeneities required of the optical components have been shown not to be present. In this paper I show that superposition images are formed in the eyes of a midwater shrimp, but in a different manner (Fig. 1b).

Live *Oplophorus spinosus* (Coutière), about 5 cm long, were obtained from trawls at a depth of 500 m in the North Atlantic during a recent cruise of the RRS Discovery (July–August 1976). The eyes are almost spherical, and when illuminated they exhibit a bright orange glow over a circular area whose diameter (905 μm) is slightly more than half the diameter of the eye itself (1,580 μm). The glow does not disappear on exposure to light, indicating the absence of occlusory distal pigment such as is found in moths[6]. After fixation for 1 h in a mixture of glutaraldehyde and formalin in cacodylate-buffered seawater, which stiffens the tissues without loss of transparency, an eye was cut in half, and the layout of the component parts is shown in Fig. 2b. As in insect superposition eyes, the rhabdom layer has a radius of curvature approximately equal to half that of the eye surface.

A superposition image was clearly visible when the half eye was illuminated by a small source (a light guide) at a distance of 10 cm. This produced a single small but somewhat diffuse spot of light which moved around the rhabdom layer in the same direction as the source—that is, the image was erect. Occluding different parts of the cornea showed that light contributing to the image entered the eye over approximately half its diameter. When two sources were used distinct images were observed, provided the sources were separated by 15° or more at the eye.
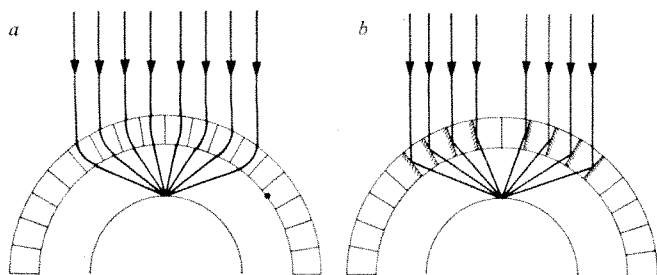
Dissection of unfixed material failed to reveal any 'crystalline' cones, although clear gelatinous material was present in the spaces in the distal pigment region, and also in the 'clear zone' between this and the rhabdoms. This is in marked contrast to the situation in, for example, euphausiid eyes where hard highly refractile cones are present as a layer beneath the cornea. The corneal facets themselves have negligible focusing properties. After fixation the gelatinous material stiffened into recognisable square-sided tapering structures (Fig. 2b) which could be easily separated from each other. Each of these consisted of a short non-pigmented outer section, a second more refractile segment extending to a depth of 150 μm from the eye surface, and lined with pigment for the proximal two-thirds of its length, and a long tapering tail whose tip lay 280–300 μm deep, in the layer of cell bodies immediately above the rhabdoms. An estimate of the refractive indices of the parts of fixed cones was made by finding cones lying with their side faces at 45° to the microscope slide, and measuring the prismatic bending of parallel rays passing through them. The distal region gave a deviation of 5.9°, corresponding to a refractive index of 1.41, and the proximal tail gave 3.3° or an index of 1.37. These values are quite in keeping with the consistency of the material, and are far too low to be compatible with image formation by refraction, using curved surfaces or lens–cylinder optics. (Euphausiid cones, treated similarly, gave a refractive index of 1.54.)

Teasing fresh material from the distal region showed that the cones are separated from each other by two kinds of pigment: a layer of reflecting material which lines the four faces of each cone and has a green specular appearance at near normal incidence, and behind this a brown pigment which fills the gap between the cones. Optically, therefore, each cone can be thought of as a silvered, four-sided prism.

This observation leads directly to the hypothesis of image formation outlined in Figs 1b and 2a. Parallel light reaching the eye is reflected at one of the faces of each cone and is bent through an angle approximately equal to twice the angle of incidence at the eye, just as in a refracting superposition eye (Fig. 1). The angle is in fact slightly greater than this, because of the 8° taper of the cones. A consequence of this is that many facets will redirect light to an approximate focus on the surface of a sphere concentric with the eye, and with slightly more than half its radius of curvature. The limiting ray will be the one that just manages to enter and emerge from a reflecting prism (left-hand ray, Fig. 2), and this makes an angle of 60° with the eye surface. This will also be the ray which emerges nearest the edge of the patch of 'glow' and it will be located approximately half way out from the centre of the eye. It can be seen from Fig. 2a that some light will tend to pass straight through the aperture of the central cones without being reflected (dashed lines). The fate of this light is uncertain: some will certainly be totally internally reflected by the walls of the tails of the cones (angles of incidence more than the critical angle, 77°) and the image is not likely to be greatly spoiled.

Two other features of the glow tend to confirm the hypothesis just given. Towards the edges of the glow patch light from individual facets is unevenly distributed, each showing a dark band which gets wider and finally occludes the facet completely. This 'venetian blind' effect is simply explained by the progressive restriction in the width of the slit of light that can enter and emerge from each reflecting prism (Fig. 2a). Secondly, the colour of the glow changes from yellow/orange in the centre (that of the bleached rhabdoms) to green at the edge of the patch. If, as seems likely, the mirror surface is due to thin-film interference[7], one would expect that all wavelengths would be reflected at high angles of incidence (near the centre), but that the light would take on progressively the interference colour of the

Fig. 1 Superposition image formation (*a*) by refraction (modified from Exner) and (*b*) by reflection.

Fig. 1 Superposition image formation (*a*) by refraction (modified from Exner) and (*b*) by reflection.

Fig. 2 Optical (*a*) and anatomical structure (*b*) of part of the eye of the shrimp *Oplophorus*. In *a* it is assumed that the only optical components in the eye are the reflector-lined faces of the cones. Solid lines show how reflected light is brought to a focus, and dashed lines indicate pencils of rays that pass through directly. The figure was constructed as accurately as possible from measurements made on fresh or recently fixed eyes. The numerals on the left show depths of each layer from the surface. *b* shows the parts of the eye as visible in a hemisected eye seen through a dissecting microscope.
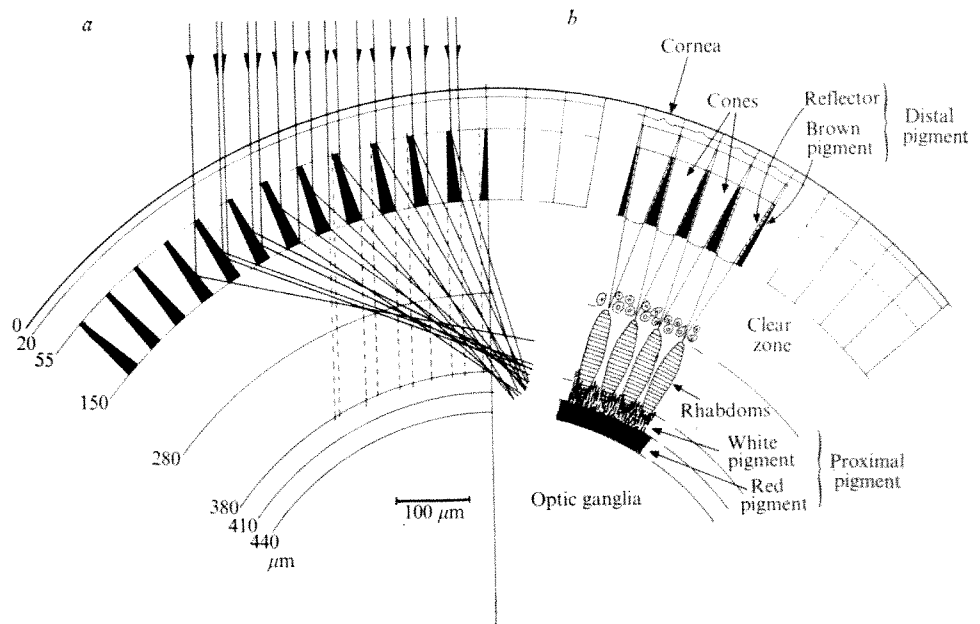
reflecting material itself at lower incident angles. Where this angle reaches 60°, at the edge, the green colour observed is similar to that of the reflecting material observed directly.

It is concluded that reflection in the cones provides an adequate explanation of the images observed, and that a refraction mechanism is not present. Since this paper was first submitted, I have learned from Professor K. Kirschfeld (Tübingen) that exactly the same optical arrangement as that given here has already been proposed independently by Klaus Vogt (Stuttgart) to account for image formation in the eyes of freshwater crayfish[8], which also exhibit a 'glow' when dark adapted[9]. His note was published last year. It also turns out that the theoretical possibility of imagery of this kind was in fact dismissed by Exner in this 1891 monograph, and the discovery of this mechanism as a practical reality properly belongs to Vogt. His paper and mine leave little doubt that this is the standard method of superposition image formation in decapod crustacea.

I thank Dr Peter Herring for making possible my participation in the Discovery cruise, Professor Eric Denton for critically reading the manuscript, and Dr Adam Locket for help with suitable fixatives. Financial assistance is acknowledged from the SRC and the Royal Society.

M. F. LAND

School of Biological Sciences,
University of Sussex,
Brighton BN1 9QG, UK

1 Loew, E. R., *Proc. R. Soc.*, B193, 31 (1976).
2 Exner, S., *Die Physiologie der facettirten Augen von Krebsen und Insekten* (Deuticke Leipzig and Wien, 1891).
3 Horridge, G. A., *The Compound Eye and Vision of Insects* (edit. by Horridge, G.A.), 255 (Clarendon, Oxford, 1975).
4 Waterman, T. H., *The Physiology of Crustacea*, II (edit. by Waterman, T. H.), I (Academic, New York and London, 1961).
5 Carricaburu, P., *The Compound Eye and Vision of Insects* (edit. by Horridge, G. A.), 236 (Clarendon, Oxford, 1975).
6 Höglund, G., *The Functional Organization of the Compound Eye* (edit. by Bernhard, C. G.), 77 (Pergamon, Oxford, 1966).
7 Land, M. F., *Prog. Biophys. molec. Biol.*, 24, 75 (1972).
8 Vogt, K., *Z. Naturforsch.*, 30c, 691 (1975).
9 Kunze, P., *Z. vergl. Physiol.*, 76, 347 (1972).

Frequency-dependent selection at two enzyme loci in *Drosophila melanogaster*

RECENT development in population genetics indicate that it is often misleading to treat selective values as constants[1-5]. These studies indicate that selective forces may be a function of genotype frequencies (frequency-dependent selection).

For example, if the rarest genotype is most fit then the rare allele will increase in frequency but will not become fixed in the population, because as it gets more common the fitness of its carriers decreases. This mechanism is attractive since it leads to a stable gene frequency equilibrium which may not involve problems of genetic load or inbreeding depression. Experimental evidence has been obtained[6,7] which supports the view that frequency dependence is involved in the maintenance of the esterase-6 (*Est-6*) and alcohol dehydrogenase (*Adh*) loci in *Drosophila melanogaster*. However, these results are subject to two major criticisms[8-11]. First, the observed frequency dependence may have resulted from poor experimental design. In the original report[6] fitness estimates were made by comparisons of adult frequencies with those expected from the gene frequencies in the previous generation. Frequency-dependent selection appears as an artefact of this method[12]. Second, since inbred lines were used in these studies problems of linkage disequilibrium may have occurred. If this view is correct then the observed selective differences were due to differences between lines, not to differences between alleles at a locus. As a consequence of these criticisms the role of frequency dependence as a force responsible for maintaining enzyme polymorphism has been disputed[9]. The experiments reported here overcome these criticisms and indicate that frequency-dependent selection should be reconsidered as an important mechanism involved in the maintenance of enzyme polymorphism.

From the Kaduna population, two pairs of lines were derived. One pair were segregating at the *Adh* locus while homozygous at the *Est-6* locus for the *F* and *S* alleles respectively whereas the other pair were homozygous at *Adh* and segregating at *Est-6*. Each line was derived independently as the progeny of 30 pairs of homozygous parents sampled from the population. The lines had been kept in population cages (as large populations) for approximately 18 months before the experiments were carried out. For each experiment 200 newly-emerged larvae (0–3 h old) were competed. They were put in a 10×4-cm *Drosophila* vial containing 10 ml of the following medium : maize meal, 26.25 g; molasses, 26.25 g; agar, 10 g; flaked brewers yeast, 2.5 g; nipagin, 2.5 g; propionic acid, 2.5 ml; water, 1,000 ml. The frequency of the *F* allele in the experimental vials was altered by changing its relative proportion in the larval mixture (for example, a vial with $P_f = 0.15$ would contain 30 *FF* homozygotes and 170 *SS* homozygotes). For the *Est-6* experiment seven frequencies of the *F* allele were used ($P_f = 0.15$; 0.25; 0.40; 0.50; 0.60; 0.75; 0.85) and for the *Adh*

Table 1 The relative viabilities of *FF* and *SS* flies when grown at different frequencies

Enzyme locus	P_f	FF input*	SS input*	FF survived (%)	SS survived (%)	Total survived (%)
	0.15	120	680	55.8	34.0	37.3
	0.25	200	600	46.5	38.0	40.1
	0.40	320	480	40.9	46.9	44.5
Est-6	0.50	400	400	38.8	43.0	40.9
	0.60	480	320	27.1	46.9	35.0
	0.75	600	200	38.0	57.0	42.8
	0.85	680	120	33.2	68.3	38.5
	0.15	120	680	65.8	45.7	48.8
Adh	0.50	400	400	43.8	35.3	39.5
	0.85	680	120	53.4	52.5	53.3

*Each result is the sum of four pooled replicates.

experiments three frequencies of the *F* allele were used ($P_f = 0.15$; 0.50; 0.85).

The genotypes of all emerging flies were determined by the methods of O'Brien and MacIntyre (*Est-6*)[13] or Day *et al.* (*Adh*)[14]. The results are shown in Table 1.

The results have been analysed by the heterogeneity χ^2 method (Table 2). The regression of the difference in survival between *FF* and *SS* flies on frequency is significant for both the *Est-6* and *Adh* experiments (*Est-6*, $P<0.005$; *Adh*, $P<0.01$). These results clearly demonstrate frequency dependence operating at both the *Est-6* and *Adh* loci. For the *Est-6* locus at *F* frequencies less than about 30% more *FF* flies survive than *SS* flies. This situation is reversed at frequencies above about 40% *F*. Here it would seem that a frequency-dependent mechanism is sufficient to maintain the polymorphism, at least in these culture conditions. This is not the case for the *Adh* polymorphism since more *FF* flies survive than *SS* flies at each frequency. The results for the *Adh* polymorphism differ from those of Kojima and Tobari[7] since these authors presented data which showed that the relative advantage of the *F* allele changed to a disadvantage as its frequency increased. My results have been confirmed by Briscoe[15] who used the same materials but different experimental methods. In a further communication (Morgan, in preparation) I will present data that indicates that the degree of frequency dependence at the *Adh* locus can be modified by growing flies on different culture media.

The results presented in this paper suggest that the role of frequency dependence in maintaining enzyme polymorphism should be examined in greater detail.

Table 2 Heterogeneity χ^2 of the differences in survival between *FF* and *SS* flies grown at different frequencies*

Est-6	d.f.	χ^2
Average difference in survival between *FF* and *SS*	1	29.66*
Linear regression of $S_f - S_s$ on P_f	1	99.00*
Departure from linearity	5	9.25†
Replicates	21	23.74†
Adh		
Average difference	1	14.86*
Linear regression	1	7.54‡
Departure from linearity	1	0.17†
Replicates	9	4.15†

For each trial, differences in survival between the two genotypes may be tested as χ^2_1, which equals $nP_f (1 - P_f) (_fS - S_s)^2/S(1 - S)$, where *n* is the total number of individuals at risk, P_f is the starting frequency of *FF* flies and S_f, S_s and S respectively are the proportions of *FF*, *SS* and all flies which survive. Over a set of trials, the sum of these values may be partitioned as in the table.
*$P < 0.005$.
†No significant difference: $P > 0.05$.
‡$P < 0.01$.

*Department of Genetics,
Monash University,
Clayton, Victoria 3168,
Australia*

1 Moment, G. B., *Science*, **136**, 1055–1056 (1962).
2 Allen, J. A., and Clarke, B., *Nature*, **220**, 501–502 (1968).
3 Levin, D. A., *Am. Nat.*, **106**, 453–460 (1972).
4 Nassar, R., Muhs, H. J., and Cook, R. D., *Evolution*, **27**, 558–564 (1973).
5 Clarke, B., in *The Role of Natural Selection in Human Evolution* (edit. by Salzano, F. M.), 187–200 (North-Holland, Amsterdam, 1975).
6 Kojima, K., and Yarbrough, K. M., *Proc. natn. Acad. Sci. U.S.A.*, **57**, 645–649 (1967).
7 Kojima, K., and Tobari, Y., *Genetics*, **61**, 201–209 (1969).
8 Yamazaki, T., *Genetics*, **67**, 579–603 (1971).
9 Lewontin, R. C., *A. Rev. Genet.*, **7**, 1–17 (1973); *The Genetic Basis of Evolutionary Change* (Columbia University Press, New York, 1974).
10 Dolan, R., thesis, Edinburgh Univ. (1974).
11 Dolan, R., and Robertson, A., *Heredity*, **35**, 311–316 (1975).
12 Prout, T., *Evolution*, **19**, 546–551 (1965).
13 O'Brien, S. J., and MacIntyre, R. J., *Am. Nat.*, **103**, 97–113 (1969).
14 Day, T. H., Hillier, P. C., and Clarke, B., *Biochem. Genet.*, **11**, 141–153 (1974).
15 Briscoe, D. A., thesis, Edinburgh Univ. (1973).

X-ray induced mutations, DNA and target theory

THAT X-ray induced, specific locus, germ-line mutation rates vary significantly in eukaryotes is generally known, but the factors that are responsible for such differences have not been assessed in detail. It was recently proposed[1] for a variety of species that mutability is closely and simply related to total DNA per genome. This approach, however, has not stood the test of critical review[2]; in fact, even within one species (mouse) there are major differences between the rates for different genes, and even for different germ cells[3–5]. I have therefore taken a different approach by considering the radiobiology of the individual gene. The analysis leads to three major conclusions. (1) Mutation rates tend to be much lower than radiation theory predicts. (2) Selection and/or repair are major factors that determine the rates. (3) The mouse 7-locus test, which provides a principal data base for the standards of human radiation hygiene, may not provide adequate overall representation of the mutability of the mammalian genome, so more research is needed.

Before considering the differences in rate between organisms, it is of interest to examine the differences between genes that are irradiated in the same germ cell of a particular organism. As an elementary step, consider the following analysis that is based on classical target theory. The fundamental premise is that bigger genes (molecules) are more likely to be hit than smaller ones and therefore should be more mutable. The calculation of target size involves the relation between observed hits and the distribution of ion pairs generated in the exposed material by the absorbed radiation, and has been applied to the study of enzymes and viruses[6,7].

In the usual germ-line mutation experiment, a parent is irradiated and the F_1 progeny are screened for the mutant phenotype. The mutation rate of a particular gene is calculated as mutations per rad per generation where per generation indicates that the observations were made on the viable progeny. But the frequency of observed hits (mutants) in these progeny is unlikely to be equal to the frequency of hits (mutations) in the parental germ cell since: (1) some hits will be repaired; (2) some hits will be carried in cells that die out or are selected against; (3) some hits will induce trivial mutations, that is, the mutation does not induce a significant change in phenotype although

provided adequate opportunity to do so. These "deflating" factors will tend to make the target-theory estimate of gene size too small.

There are also "inflating" factors that tend to make the estimate too large. First, in addition to direct hits, free radicals generated in nearby molecules (for example water) may mutate the target gene. Secondly, since "target" is an operational term, its volume might in fact represent the aggregate volume of several separate genes, any one of which when hit can directly or indirectly induce the mutant phenotype.

It is clear that the estimate of gene size by target theory will probably be incorrect, depending on the magnitude and balance of a number of factors. Nonetheless, such an incorrect estimate can be useful for purposes of comparison. As the standard for such comparison, we use the molecular weight of a typical gene. Since each codon has a molecular weight of about 2,000, a gene that determines a protein of 200 amino acids has a molecular weight of about 400,000.

I applied target theory to the data of the three tests with mouse spermatogonia (250-kV X rays, 60–90 rad min^{-1}), using the method outlined by Lea (chapter 3, fig. 9 of ref. 6). The method is adequate for the comparisons to be made here, although it does not provide a full treatment of the problem[8–10].

The 7-locus test used one cartilage and six coat-colour genes. For a dose of 300 rad, the mean probability K of inducing a mutation was 2.7×10^{-7} per gene per rad per generation[4]. Greater doses or lower dose rates reduced the probability. The mean dose per mutation (Lea's 37% dose) is therefore taken as $1/K = 3.7 \times 10^{6}$ rad, for which the associated molecular weight is 360,000 (180 codons). The close agreement of this estimate with that based on the model is presumably fortuitous. Nonetheless, the fact that there is numerical agreement rather than gross disagreement gives perspective on the balance of deflating and inflating factors.

In a very much smaller set of experiments with an analogous "6-locus test", the value of K was about one-third that of the 7-locus test[11].

In the case of the H-test (30 class I histocompatibility genes)[5], the mean value of K was 1/60 that of the 7-locus test, and the associated molecular weight is therefore about 5,000—an underestimate equivalent to 2.5 codons. The underestimate is informative, however. It indicates that the set of deflating factors is much more important than the inflating ones.

With two such disparate results for the mouse as a point of departure, their generality becomes of interest—will various genes in diverse species behave like the 7-locus group or the H-group? The literature was recently reviewed[2] for X-ray induced, intralocus, forward mutation rates. Determinations were found for *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae* (yeast), *Neurospora crassa* (mould), *Drosophila melanogaster* (fruit fly), *Arabadopsis thaliana* (flowering plant). *Zea mays* (maize) and *Hordeum vulgare* (barley). From the summary, I obtained the distribution of 16 values of K for different genes and different groups of genes. The highest value was 4.7×10^{-8}, equivalent to 25 codons. Eleven fell between 3×10^{-8} and 4×10^{-9}, with a mean of 1.5×10^{-8} (7 codons), and four were below 3×10^{-9} (1 codon).

It seems that the genes or groups of genes tested so far (outside the mouse) tend to react like the mouse H-group rather than the 7-locus group, and that the factors that deflate the mutation rate (per gene per generation) are much more effective than those that inflate it. To what extent this conclusion will continue to be generally true for all types of genes, and especially for mutations other than forward ones, remains to be seen. Meanwhile, the analysis provides some specific orientation towards the general problem and emphasises the need for much more

research, both for genetics *per se* and for radiation hygiene, which relies heavily on the 7-locus test as a data base for setting standards.

HENRY I. KOHN

Department of Radiation Therapy,
Harvard Medical School, and
*   Shields Warren Radiation Laboratory of the*
*   New England Deaconess Hospital,*
Boston, Massachusetts 02115

1 Abrahamson, S., Bender, M. A., Conger, A. D., and Wolff, S., *Nature*, 245, 460–461 (1975).
2 Schalet, A. P., and Sankaranarayanan, K., *Mutat. Res.*, 35, 341–370 (1976).
3 Russell, W. L., in *Advances in Radiation Research Biology and Medicine* (edit. by Duplan, J. F. and Chapiro, A.), 323–334 (Gordon and Breach, New York, 1973).
4 Russell, W. L., in *Repair from Genetic Radiation Damage*, (edit. by Sobels, F. H.), 205–217 (Pergamon, Oxford, 1963).
5 Kohn, H. I., and Melvold, R. W., *Nature*, 259, 209–210 (1976).
6 Lea, D. E., *Actions of Radiations on Living Cells* (Cambridge University Press, Cambridge, 1946).
7 Pollard, E. C., Guild, W. R., Hutchinson, F., and Setlow, R. B., *Progress in Biophysics*, 5 (edit. by Butler, J. A. V., and Randall, J. T.), 72–108 (Academic, New York, 1955).
8 Hutchinson, F., *Science*, 134, 533–538 (1961).
9 Dertinger, H., and Jung, H., *Molecular Radiation Biology* (Springer, New York, 1970).
10 Augustine, L. G., Brustad, T., and Mason, R., *Adv. Radiat. Biol.*, 1 (edit. by Augenstein, L. G., Mason, R., and Quastler, H.) 227–266 (Academic, New York, 1964).
11 Lyon, M. F., and Morris, T., *Mutat. Res.*, 8, 191–198 (1969).

Culture of human malaria parasites *Plasmodium falciparum*

MALARIA is a major cause of morbidity and mortality, claiming an estimated one million lives a year in Africa alone[1]. *Plasmodium falciparum*, the most important agent of human malaria, has not been maintained for more than 2–6 d *in vitro*[2–10]. The development of a vaccine for malaria depends on suitable culture methods for the production of relevant antigens[1]. We describe here a new culture method for *P. falciparum*, which involves an inoculum of cryopreserved parasites and make possible an analysis of merozoite–erythrocyte interactions, growth for more than 3 weeks *in vitro* and the collection of merozoites. Merozoites are of particular interest because vaccination with the merozoites of *P. knowlesi* protects rhesus monkeys from infection with this species of malaria parasite[11].

Parasites for culture were obtained from a splenectomised chimpanzee (*Pan troglodytes*) inoculated with a chimpanzee-adapted isolate of the Malayan Camp strain of *P. falciparum*[12]. Nine days after inoculation 52% of its erythrocytes were parasitised with ring forms and 6% with schizonts. A sample (50 ml) of this infected blood was collected in 500 U of heparin and cryopreserved in 60 aliquots by the method of Diggs *et al.*[3].

Growth of *P. falciparum* in erythrocytes from five species was tested. Blood was drawn in heparin from human volunteers, owl monkeys (*Aotus trivirgatus griseimembra*), rhesus monkeys (*Macaca mulatta*), Hartley guinea pigs (*Cavia porcellus*) and also from the chimpanzee which had been infected (and then cured with mefloquine[13]). These erythrocytes were separated from plasma and leukocytes, mixed with an inoculum of thawed parasitised chimpanzee erythrocytes, and the mixture was cultured at a low oxygen tension in modified tissue culture medium 199. Samples were taken at the times indicated, thin smears were made and stained with Giemsa and 1,000 erythrocytes were examined for parasites. A representative experiment is presented in Table 1. Erythrocytes from humans and the chimpanzee, both of which can be infected *in vivo* with the Camp strain of *P. falciparum*, were parasitised *in vitro*. Erythrocytes from rhesus monkeys and guinea pigs, which cannot be infected *in vivo*, were not invaded to a significant extent *in vitro*. Other experiments provided evidence that erythrocytes from owl monkeys (which can be infected *in vivo*) support growth *in vitro*; too few experiments have

Table 1 Malaria grown in four species of erythrocytes

Test erythrocytes	Trophozoites at 24 h	Trophozoites at 72 h	Ratio of 72 h to 24 h
Human	3.7±1.5	23.0±3.1	6.2
Chimpanzee	5.3±0.9	33.3±1.7	6.3
Rhesus	3.0±1.2	1.3±0.3	0.4
Guinea pig	5.5±2.5	1.8±0.5	0.3

Numbers of trophozoites are given per 1,000 erythrocytes (mean ± s.e. of triplicate cultures).

Test erythrocytes and thawed[3] chimpanzee erythrocytes (50% parasitised with ring forms; schizonts did not survive thawing) were mixed at a ratio of 25 : 1 and cultured at a final concentration of 5×10^7 per ml in an atmosphere of 3% CO_2, 6.6% O_2, balance nitrogen at 37 °C (3 ml per 5-cm^2 flask, or 0.2 ml per microtitre well). Culture medium 199 (with Earle's modified salts) was enriched with glucose (2 mg ml^{-1}), 2 mM glutamine, 3×10^{-5} M 2-mercaptoethanol, (±)-α-tocopherol (emulsified) (30 µg ml^{-1}), gentamycin (25 µg ml^{-1}), 10% heat-inactivated foetal bovine serum, and 10 mM *N*-Tris-(hydroxymethyl)methyl-2-aminoethanesulphonate (TES) buffer. (One part TES acid to 1.4 parts TES sodium salt gives a *p*H of 7.35 at 37 °C). Leukocytes and plasma had been removed from the erythrocytes by passage through a column of cellulose powder[14] followed by washing the cells several times in media.

been done to comment on the finding[9] that they support growth of *P. falciparum* less well than do human erythrocytes.

In these short term cultures the parasites grow first in the erythrocytes of the chimpanzee inoculum (ring forms predominated at 0 h, more mature trophozoites at 24 h and schizonts at 42 h); then at 48 h merozoites invaded new erythrocytes (young ring forms were also seen at this time in the new erythrocytes). By 72 h the parasite had grown to the trophozoite stage again. At both 48 and 72 h the cultures with human erythrocytes had 15 to 50 times more parasites than cultures with rhesus erythrocytes. Small numbers of degenerating parasites were seen in all cultures. The viability of the cryopreserved parasite inoculum varied from 20 to 50% as judged by the numbers of maturing trophozoites after 24 h of culture.

Long term culture was also possible. Schizonts were concentrated by centrifugation every 2 or 4 d and reinoculated into a culture with new human erythrocytes. The parasite exhibited some asynchrony of growth: on day 22 of culture

Fig. 1 Malaria parasites after 22 d in culture. *a*, Merozoites invading erythrocytes (top right), schizonts (top left and middle), trophozoites (bottom middle); *b*, ring forms; *c*, a gametocyte. Parasites were cultured essentially as described in Table 1, except that every 2 or 4 d the erythrocytes were centrifuged at 400*g*, resuspended in a small volume of medium, centrifuged in capillary tubes at 1,300*g* for 2 min, and the capillaries were scored and broken to collect the upper brown layer of schizonts (usually 3×10^7 cells) which was added to a new 15-ml subculture containing 3×10^8 human erythrocytes in a 22-cm^2 flask. The medium was usually changed every 48 h with the subculture. After the first 4 d of culture, an extra 15 mM TES buffer was added, and 20% autologous human serum replaced the foetal bovine serum in the medium.
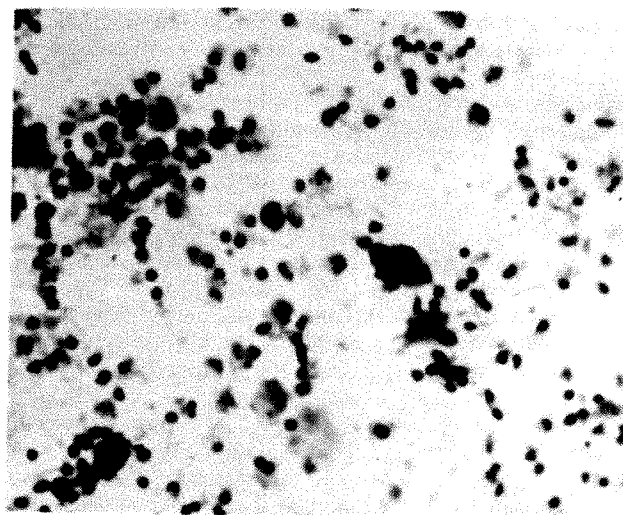
Fig. 2 Thawed parasitised chimpanzee erythrocytes (90% parasitised after a centrifugation step in capillary tubes) were cultured without added erythrocytes, but otherwise essentially as described in Table 1. After 44 h, the culture was centrifuged at 400*g* for 8 min. The supernatant was then taken and centrifuged at 1,000*g* for 10 min and the small grey pellet was resuspended, smeared and stained with Giemsa.

schizonts, invading merozoites, ring forms, trophozoites and what appeared to be immature gametocytes were seen (Fig. 1). The total number of schizonts recovered from each subculture was approximately equal to the number added 2 d previously, except on day 22 when twice as many schizonts were recovered. This increased yield followed an increase in the frequency of changing the medium from every other day to every day. A significant metabolic production of acid after 1 d in culture was indicated by the yellow colour of the phenol red in the medium. The addition of a flow system to continually replenish the medium should result in better yields. Further manipulation of culture conditions might also result in the production of mature gametocytes which could be used to study the sexual cycle *in vitro*.

Merozoites were obtained by culturing parasitised erythrocytes to the schizont stage, centrifuging at low speed to remove the unruptured schizonts, and then centrifuging at high speed to pellet the free merozoites (Fig. 2). These *P. falciparum* merozoites have not yet been examined for viability, but are expected to have a lifespan measured in minutes[15].

Experiments to evaluate the importance of the various components of the culture medium are incomplete. Others have reported improved growth of malaria parasites with the use of TES as a buffer[7], the removal of leukocytes[16] and the use of a low oxygen tension[17,18].

As expected from work with *P. knowlesi*[19-21], the growth of *P. falciparum* in various mammalian erythrocytes *in vitro* correlated with the susceptibility of the erythrocyte donors *in vivo*. This erythrocyte specificity apparently resides in the interactions between the erythrocyte and merozoite surfaces, since very few merozoites adherent to erythrocytes were noted at 48 h in cultures containing erythrocytes from non-susceptible species. The few interactions that were noted probably occurred with residual chimpanzee erythrocytes from the inoculum. This culture system is being used to study *P. falciparum* interactions with various human blood group erythrocytes with and without enzyme treatments (L. H. Miller *et al.*, manuscript in preparation), as reported for *P. knowlesi*[20-23].

Further improvements in the culture of *P. falciparum*, in particular the production of large numbers of merozoites, should facilitate biochemical and immunological studies that may lead to an effective vaccine for malaria.

While this manuscript was in preparation, Trager and

Jensen reported the continuous culture of *P. falciparum*[24].

J. David Haynes
Carter L. Diggs
Fred A. Hines
Robert E. Desjardins

Departments of Immunology,
*   Parasitic Diseases and Pharmacology,*
Walter Reed Army Institute of Research,
Washington, DC 20012

1 *Developments in Malaria Immunology, Tech. Rep. Ser. Wld Hlth Org.*, **579** (1975).
2 Diggs, C., *et al.*, *J. Parasit.*, **57**, 187–188 (1971).
3 Diggs, C., Joseph, K., Flemings, B., Snodgrass, R., and Hines, F., *Am. J. trop. Med. Hyg.*, **24**, 760–763 (1975).
4 Mitchell, G. H., Butcher, G. A., Voller, A., and Cohen, S., *Parasitology*, **72**, 149–162 (1976).
5 Pavanand, K., Permpanich, B., Chuanak, N., and Sookto, P., *J. Parasit.*, **60**, 537–539 (1974).
6 Phillips, R. S., Trigg, P. I., Scott-Finnigan, T. J., and Bartholomew, R. K., *Parasitology*, **65**, 525–535 (1972).
7 Siddiqui, W. A., and Schnell, J. V., *J. Parasit.*, **59**, 516–519 (1973).
8 Trager, W., *J. Protozool.*, **18**, 239–242 (1971).
9 Trigg, P. I., *Parasitology*, **71**, 433–436 (1975).
10 Wilson, R. J. M., and Bartholomew, R. K., *Parasitology*, **71**, 183–192 (1975).
11 Mitchell, G. H., Butcher, G. A., and Cohen, S., *Nature*, **252**, 311–313 (1974); *Immunology*, **29**, 397–407 (1975).
12 Hickman, R. L., *Milit. Med.*, **134**, 741–746 (1969).
13 Trenholme, G. M., *et al.*, *Science*, **190**, 792–794 (1975).
14 Grant, P. T., and Fulton, J. D., *Biochem. J.*, **66**, 242–250 (1957).
15 Dennis, E. D., Mitchell, G. H., Butcher, G. A., and Cohen, S., *Parasitology*, **71**, 475–481 (1975).
16 Williams, S. G., and Richards, W. H. G., *Ann. trop. Med. Parasit.*, **67**, 169–177 (1973).
17 Cohen, S., Butcher, G. A., and Crandall, R. B., *Nature*, **233**, 368–371 (1969).
18 Tiner, J. D., *Proc. helminth. Soc. Wash.*, **39**, 211–219 (1972).
19 Butcher, G. A., Mitchell, G. H., and Cohen, S., *Nature*, **244**, 40–42 (1973).
20 Miller, L. H., Mason, S. J., Dvorak, J. A., McGinniss, M. H., and Rothman, I. K., *Science*, **189**, 561–563 (1975).
21 Miller, L. H., Mason, S. J., Clyde, D. F., and McGinniss, M. H., *New Engl. J. Med.*, **295**, 302–304 (1976).
22 Miller, L. H., Dvorak, J. A., Shiroishi, T., and Durocher, J. R., *J. exp. Med.*, **138**, 1597–1601 (1973).
23 Miller, L. H., Shiroishi, T., Dvorak, J. A., Durocher, J. R., and Schrier, B. K., *J. molec. Med.*, **1**, 55–63 (1975).
24 Trager, W., and Jensen, J. B., *Science*, **193**, 673–675 (1976).

Direct estimation of frequency of cytotoxic T lymphocytes by a modified plaque assay

Thymus-derived lymphocytes (T) have been shown to have a major role in cell-mediated immunity against foreign transplants and syngeneic tumours[1,2]. *In vitro*, several assays have been developed to measure thymus-dependent cell-mediated immunity, making use of the cytotoxic potential of these effector cells[3]. Precise estimates of the number of cytotoxic T lymphocytes (CTL), however, are not possible by available methods. We report here a new method to quantify the CTL present in a mixed cell population. The method is based on the ability of CTL to kill specifically adjacent target cells fixed on monolayers and to form zones of lysis or plaques.

Monolayers of target cells were prepared on poly-L-lysine-coated Falcon plastic Petri dishes as described before[4]. The effector lymphocytes were generated after either *in vivo* sensitisation or *in vitro* sensitisation against tumour allografts. Known numbers of effector lymphocytes were added on the target monolayers and the plates were incubated at 37 °C in an atmosphere of 5% CO_2. After incubation plates were washed with minimum Eagle's medium (MEM), dead tumour cells were stained with eosin, and the plates were washed again with MEM. Zones of dead target cells or plaques were counted under the microscope. A cluster of more than four dead target cells was considered a plaque (Fig. 1).

Several representative experiments are shown in Table 1. Both the *in vivo* and *in vitro* sensitised alloimmune lymphocytes (BALB/c anti-EL-4) form plaques on EL-4 target-cell
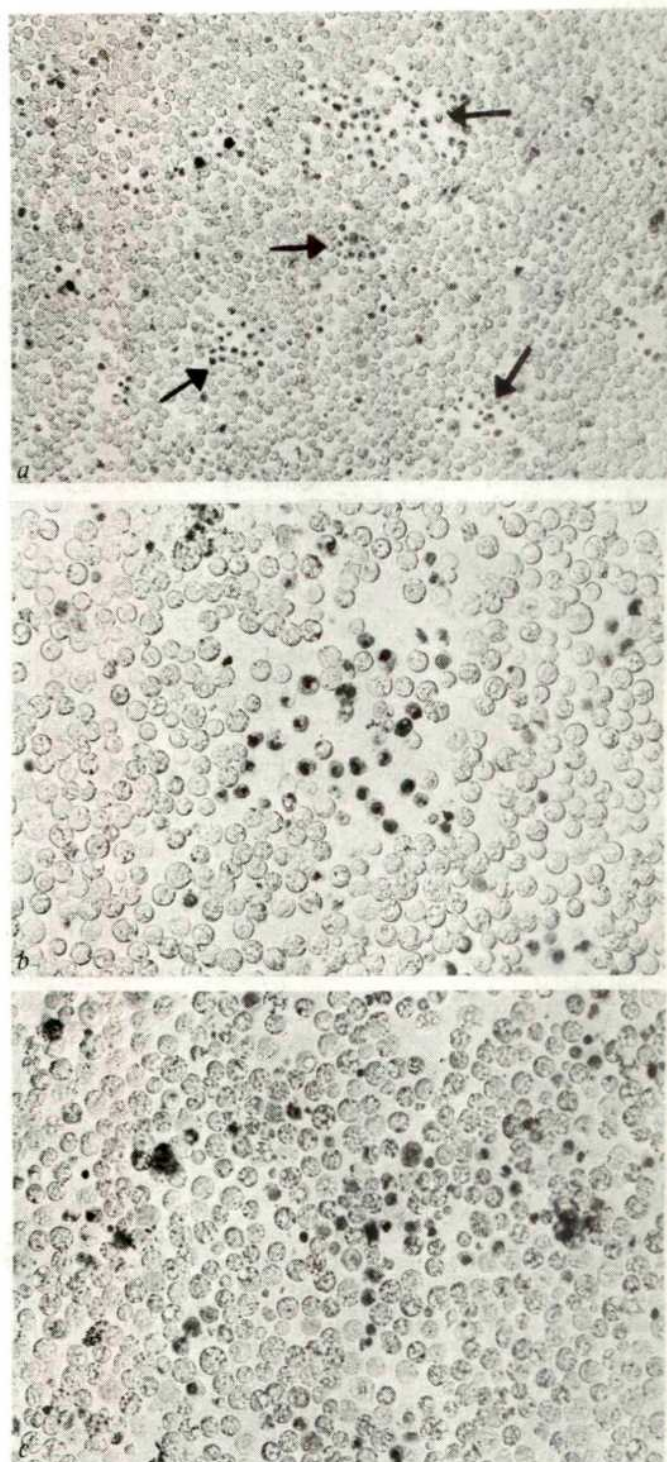
Fig. 1 A monolayer prepared with EL-4 tumour cells and plaques developed by *in vitro*-generated anti-EL-4 cytotoxic lymphocytes. *a*, Arrows point to plaques ($\times 125$); *b*, a plaque ($\times 250$); *c*, a monolayer background lacking plaques ($\times 250$).

monolayers. The specificity of plaque formation was indicated by the absence of plaques when unrelated target-cell monolayers (G-35) were used. Furthermore, normal lymphocytes or lymphocytes cultured alone did not form plaques. A linear relationship was usually obtained between the number of effector cells plated and the number of plaques observed, indicating that each plaque was presumably formed by one CTL. As expected, the *in vitro*-generated secondary response yielded many more plaques when compared with the *in vivo* immune response (experiments 1 and 4, Table 1). Treatment of effector lymphocytes with specific anti-thymocyte serum and complement abrogated completely

both the formation of plaques and the cell-mediated cyto-toxic activity obtained by the ^{51}Cr-release assay (experiment 5, Table 1). Neither serum nor complement alone had an adverse effect on plaque formation. These results demonstrated that alloimmune lymphocytes form specific plaques on corresponding target cell monolayers and that the effector cells are thymus-derived lymphocytes.

The relationship between plaque formation and the release of ^{51}Cr from target cells as measures of cell-mediated cytotoxicity was investigated. In all cases, when effector populations produced plaques, they also showed cytotoxicity by ^{51}Cr release (Table 2). In several experiments, effector cells with strong cytotoxic activity yielded a large number of cytolytic plaques. That the same population of effector cells exerted these two properties, is corroborated by the findings that reduction of cytotoxic activity by absorption on specific target cell monolayers was paralleled by a marked reduction in the number of plaques (experiment 1, Table 2). Experiments with *in vitro* sensitisation of C57BL/6 splenocytes against syngeneic EL-4 leukaemia cells have

Table 1 Frequency of CTL in alloimmune lymphocyte preparations

Experiment no.	Effector cells Sensitisation	Specificity	No. of cells plated	Monolayer	Plaque-forming units Observed	Per 10⁶	%
1	*In vitro*	Anti-EL-4	3.1×10^5	EL-4	2,380	7,850	0.79
	In vitro	Anti-EL-4	1.5×10^5	EL-4	935	6,171	0.62
	In vitro	Anti-EL-4	3.1×10^5	G-35	0	0	0
	In vitro	Anti-EL-4	1.5×10^5	G-35	0	0	0
	In vivo	Anti-EL-4 PEC	3.1×10^5	EL-4	80	264	0.03
	In vivo	Anti-EL-4 PEC	1.5×10^5	EL-4	35	245	0.03
	In vivo	Anti-EL-4 PEC	3.1×10^5	G-35	0	0	0
	In vivo	Anti-G-35 PEC	3.1×10^5	G-35	70	210	0.02
2	*In vitro*	Anti-EL-4	3.1×10^5	EL-4	2,800	9,240	0.92
	In vitro	Anti-EL-4	1.5×10^5	EL-4	1,210	8,680	0.87
	In vitro	Anti-EL-4	7.5×10^4	EL-4	600	8,400	0.84
	Lymphocytes alone		3.1×10^5	EL-4	0	0	0
	In vitro	Anti-EL-4	3.1×10^5	G-35	0	0	0
3	*In vitro*	Anti-EL-4	3.1×10^5	EL-4	(4,410; 3,780; 4,340)*	13,778	1.38
	In vitro	Anti-EL-4	1.5×10^5	EL-4	(1,990; 2,275; 2,030)	14,555	1.46
	In vitro	Lymphocytes alone	3.1×10^5	EL-4	0	0	0
4	*In vivo* (2 h)	Anti-EL-4 PEC	3.1×10^5	EL-4	210	693	0.07
			1.5×10^5	EL-4	85	561	0.06
		Normal PEC	7.5×10^4	EL-4	70	924	0.09
			3.1×10^5	EL-4	0	0	0
	In vivo (4 h)	Anti-EL-4 PEC	3.1×10^5	EL-4	840	2,772	0.28
			1.5×10^5	EL-4	210	2,793	0.28
		Normal PEC	3.1×10^5	EL-4	0	0	0
5	*In vitro*	Anti-EL-4	6.25×10^5	EL-4	2,100	3,360	0.34
	In vitro	Anti-EL-4	3.1×10^5	EL-4	1,890	4,515	0.45
	Lymphocytes alone		6.25×10^5	EL-4	0	0	0
	In vitro	Anti-EL-4+ Anti T+C*	6.25×10^5	EL-4	0	0	0
			3.1×10^5	EL-4	0	0	0
	In vitro	Anti-EL-4	6.2×10^5	G-35	0	0	0
6	*In vivo*	Anti-EL-4 PEC	2.5×10^6	EL-4	840	336	0.034
	Normal PEC		2.5×10^6	EL-4	0	0	0
	In vivo	Anti-EL-4 PEC	2.5×10^6	C57PM	750	300	0.030
	Normal PEC		2.5×10^6	C57PM	5	2	0

Effector cells (alloimmune lymphocytes) were prepared after *in vivo* and *in vitro* immunisation. Cytotoxic cells were generated *in vivo* by intraperitoneal inoculation of 2×10^7 EL-4 (H-2^b) into BALB/c (H-2^d) mice for anti-EL-4 activity, or 2×10^7 G-35 (H-2^d) syngeneic in BALB/c into C57BL/6 (H-2^b) for anti-G-35 activity. The mice were killed 9–10 d later and the peritoneal fluid exudate was aspirated by a Pasteur pipette after several washes of the cavity. The cells were washed, suspended in MEM at 10^7 per ml and samples of 1 ml were plated on plastic tissue culture dishes. The plates were incubated at 37 °C for 1 h and the non-adherent cells were recovered, counted for viability by Trypan blue dye exclusion and suspended in RPMI or MEM at desired concentrations. Lymphocytes were generated *in vitro* as follows. Mice were primed with 2×10^7 tumour cells and 40–60 d later they were killed and their spleens were removed in sterile conditions. The cells were teased apart in balanced salt solution (BSS), filtered through a nylon screen and washed once. Cells were then treated with Gey's solution to lyse red blood cells, washed twice with BSS and resuspended to 10^7 per ml viable cells in RPMI 1640 containing 10% foetal calf serum (FCS). The stimulating tumour cells were prepared as above for responding cells and treated with 30 μg mitomycin C per 10^7 cells. The cells were incubated at 37 °C for 40 min and washed four times with BSS and resuspended to 5×10^6 per ml viable cell in RPMI 1640 containing 10% FCS. Sensitisation was done in 50-ml Falcon conical tubes in which 3 ml of the responder cells (30×10^6 per ml) and 1 ml of stimulator (5×10^6 cells) were mixed and the total volume was brought up to 10 ml. The tubes were incubated for 3–5 d at 37 °C in a 5% CO_2 incubator. At the end of incubation, dead cells were removed using Ficoll–Isopaque. The cell preparation was then brought to the desired concentration in RPMI 1640 + 10% FCS and used for plaque assays or for ^{51}Cr-release cell-mediated cytotoxicity. Monolayers were prepared on 35×10-mm Falcon plastic dishes (no. 3001) coated with poly-L-lysine. One ml of poly-L-lysine (50 μg ml^{-1}, molecular weight 85,000, Sigma, Mi) in phosphate-buffered saline (PBS) (Dulbecco's PBS + $CaCl_2$ + $MgCl_2$) was added to each plate and left at room temperature overnight. Plates were washed with PBS before use. Target cells used to prepare the monolayers were either tumour cells or peritoneal macrophages (PM). Tumour cells (EL-4 and G-35) were maintained in ascites form *in vivo*. The ascites fluid containing tumour cells was washed and tumour cells were suspended in PBS at 10^7 per ml. Target cells (1 ml) were added to each plate, which were then left at room temperature for 10 min, centrifuged at 200g for 5 min at 20 °C and washed three or four times with PBS. In most experiments, 1 ml of 0.2% Trypan blue in saline was added to each plate, left at room temperature for 3–5 min, washed with PBS, and monolayers were examined for viability under the microscope. PMs were derived from mice inoculated with 3 ml of thioglycollate 3 d before collection. Peritoneal cells were prepared as described above for tumour cells. Plaque-forming units were estimated as follows. Effector lymphocytes were added at desired concentrations in a total of 1 ml per plate. Plates were incubated at 37 °C in an atmosphere 5% CO_2 for 2–4 h. Plates were washed with 2 ml MEM, and stained with 1 ml of 1% eosin dye for 3–5 min. The plates were washed with PBS and examined under the microscope. In several experiments, monolayers were fixed by treating the plate with 0.2% formaldehyde overnight at 5 °C and then washed with PBS. Plaques were scored under the microscope. Manual interplay with the condenser facilitated location of plaques. In general, the plates were placed on a microscope slide with a square drawn calibrated for 1/5 the surface area of the Petri dish. The total number of plaques per square was routinely counted and the number of plaques per 10^6 lymphocytes was thus estimated.

*In this experiment each effector cell concentration was tested in triplicate.

†Treatment with specific anti-thymocyte serum and C' (ref. 11) was as follows: 0.3 ml of 1:10 dilution of specific rabbit anti-thymocyte serum and 0.3 ml of 1:10 normal rabbit serum (absorbed with tumour cells) as source of complement was added to 10^6 lymphocytes; the mixture was incubated at 37 °C for 45 min and the cells were washed with MEM and resuspended at the original concentration.

Table 2 Correlation between plaque formation and cell-mediated cytotoxicity

Experiment no.	Effector cells	No. of cells plated	Monolayer	Plaque-forming units Observed	%	% CMC to ^{51}Cr EL-4
1	BALB/c anti-EL-4	1×10^5	EL-4	1,650	1.65	97.0
	BALB/c anti-EL-4 absorbed on EL-4 monolayer*	1×10^5	EL-4	225	0.22 (87)†	49.1 (49)
	BALB/c anti-EL-4 absorbed on G-35 monolayer	1×10^5	EL-4	1,540	1.54	83.6
	BALB/c anti-EL-4	1×10^5	G-35	0	0	0
2	BALB/c anti-EL-4	1×10^5	EL-4	420	0.42	58.7
	Normal BALB/c spleen	5×10^5	EL-4	15	0.003	0

Effector cells were generated *in vitro* in one way mixed lymphocyte–tumour cultures. 5×10^6 BALB/c splenocytes were mixed with 2.5×10^6 mitomycin C ($50 \ \mu g \ ml^{-1}$)-treated EL-4 cells in tubes (125×16 mm; Falcon no. 3033) in RPMI culture medium supplemented with 10% foetal calf serum. Six days later cultures were collected and the cells were separated on Ficoll–Hypaque to remove dead cells. Cell-mediated cytotoxicity (CMC) with ^{51}Cr target cells was assayed as described previously[4] at an effector–target cell ratio of 10:1.
*Absorption on monolayers was done as described before[4].
†Numbers in parentheses represent percentage reduction.

produced 5–10 times less plaque-forming units (0.04–0.12%) than allogeneic sensitisation. The ^{51}Cr release in the syngeneic sensitisation was weaker than the allogeneic system.

Several parameters have been examined in an effort to generalise the applicability of the CTL plaque assay and also to optimise the sensitivity of the assay. For example, we found that monolayers can be prepared with target cells other than tumour cells, such as thymocytes or thioglycollate-induced peritoneal macrophages (experiment 6, Table 1). The plates containing the plaques can be fixed with 0.2% formaldehyde and kept at 4 °C without any significant change in the number of plaques. Spontaneous lysis of target cells on the monolayer interferes with plaque reading. Plaques were assayed only when the monolayer target cell showed >95% viability. To distinguish between dead cells present at the initiation of incubation from those which occurred after incubation, a double staining technique was used. The target cell monolayers were first stained with Trypan blue and after plating of the lymphocytes, the monolayers were stained with eosin. The time of incubation of the assay varied from 2 to 5 h depending on the source of effector cells used. *In vitro* sensitised cells required less than 2 h, whereas the *in vivo* sensitised lymphocytes required 4 h (experiment 4, Table 1). Longer incubation periods resulted in a mixture of clear and stained plaques, and these were difficult to quantify and reproduce.

The percentage of CTL in mixed alloimmune preparations as obtained by the plaque assay ranged from 0.05 to 2% of the total plated on the monolayer. These percentages may be underestimates due to the limitations of the assay conditions. For example, (1) CTL may be heterogeneous in their state of differentiation and may express different degrees of cytotoxic activity. In such cases, only CTL that can kill more than four target cells will be scored while CTL which will kill less than four target cells will be missed in our assay. (2) The plating efficiency is not high, that is, CTL with high affinity to target cells may be preferentially selected, and (3) lysis of target cells may not be revealed by dye-exclusion techniques.

Although our plaque assay measured directly the number of CTL, indirect methods have been reported which estimate the number of CTL by their ability to bind to target cells and form conjugates. For example, with alloimmune peritoneal exudate cells (PEC), the number of CTL were estimated to be >6% by Martz[5] and up to 35% by Berke *et al.*[6].

It is difficult to ascertain from our studies the precise number of precursors of CTL present in a non-immune lymphoid cell population. Such numbers, however, have been reported for precursors which mediate mixed lymphocyte culture and graft *versus* host reactions. It was reported that 1–3% responsive cells became stimulated to proliferate in the mixed lymphocyte interaction[7], 1–2% induced local graft *versus* host reaction[8], 6% of parental T cells bear surface

receptors which share idiotypic determinants with IgG fractions of alloantibody specific for a particular major histocompatibility complex (MHC) haplotype[9], and 4.5–6% participate in a systemic graft *versus* host reaction in the rat[10]. If CTLs arise by expansion of the precursor pool after antigenic stimulation then the percentage of uncommitted precursor CTL for one MHC will be less than 1%. This percentage agrees with that postulated by Ford *et al.*[10].

In conclusion, the plaque assay described here was shown to enumerate specific cytotoxic T lymphocytes present in a mixed cell population. This modified plaque assay should be helpful in elucidating several aspects of cell-mediated immunity such as cytotoxic potential, affinity, heterogeneity, specificity, and morphological characteristics of effector T lymphocytes once isolated.

BENJAMIN BONAVIDA
BARBARA IKEJIRI

Department of Microbiology and Immunology,
School of Medicine,
University of California,
Los Angeles, California 90024

ELI KEDAR

Lautenberg Center for General and
* Tumor Immunology,*
Hebrew University,
Hadassah Medical School,
Jerusalem, Israel

1 Cerottini, J.-C., and Brunner, K. T., *Adv. Immun.*, **18**, 67–132 (1974).
2 Herberman, R. B., *et al.*, *Prog. Immun.*, **3**, 285–295 (1974).
3 Bloom, B. R., and Glade, P. R., *In Vitro Methods in Cell-Mediated Immunity* (Academic, New York, 1971).
4 Kedar, E., Ortiz de Landazuri, M., Bonavida, B., and Fahey, J. L., *J. Immun. Meth.*, **5**, 97–102 (1974).
5 Martz, E., *J. Immun.*, **115**, 261–267 (1975).
6 Berke, G., Gabeson, D., and Feldman, *Eur. J. Immun.*, **5**, 813–181 (1975).
7 Wilson, D. B., Blyth, J. L., and Nowell, P. C., *J. exp. Med.*, **128**, 1157–1181 (1968).
8 Nisbet, N. W., Simonsen, M., and Zaleski, M., *J. exp. Med.*, **129**, 459–467 (1969).
9 Binz, H., and Wigzell, H., *J. exp. Med.*, **142**, 1218–1230 (1975).
10 Ford, W. L., Simmonds, S. F., and Atkins, R. C., *J. exp. Med.*, **141**, 681–696 (1975).
11 Bonavida, B., *J. Immun.*, **112**, 1308–1321 (1974).

In vitro and *in vivo* radiosensitivity of human tumour cells obtained from a pancreatic carcinoma xenograft

HUMAN tumours show a wide range of clinical response to radiotherapy, but the extent to which this reflects differences in intrinsic cellular radiosensitivity, extent of tumour cell hypoxia, or host reaction against the tumour is unknown. For tumours in laboratory animals, various assays for clonogenic cells are available. These have shown that the range

of radiosensitivity of aerobic cells is quite narrow[1,2] and that the hypoxic fraction of clonogenic cells within tumours of palpable dimensions is often in the range 10–35%. Until now the lack of satisfactory assay for clonogenic cells has prevented similar studies on human solid tumours treated *in vivo*. We have developed two assays for clonogenic cells from a human metastatic pancreatic carcinoma propagated as a xenograft in immune-suppressed CBA/lac mice. This xenograft was one of a series established by Pickard *et al.*[3].

In the first assay, tumour-cell suspensions were incubated in Ham's medium with the addition of 15% calf serum, 0.3% agar and rat red blood cells in 17-mm diameter test tubes using a double layer technique. Colonies of 50 cells or more were counted at 28 d and the plating efficiency was 20–40%. In the second assay, tumour-cell suspensions in medium, serum and agar were grown in diffusion chambers implanted into the peritoneal cavity of C57BL mice, previously treated with whole-body radiation. Colonies, growing slightly more quickly than in the first assay, were counted at about 18 d and the plating efficiency was 10–20%. Detailed descriptions of the experimental methods are being prepared for publication. Chromosomal analysis of the cells after 16 d growth in culture verified a human karyotype, with aneuploidy ranging from 42 to 68 chromosomes and a mode of 62.

Tumours measuring 5–8 mm in diameter were grown in the leg muscles of immune-suppressed mice and were irradiated *in vivo* using ^{60}Co γ rays. The animals were killed and the tumours removed either immediately after irradiation or 18 h later, and cell suspensions were prepared for assay. Acute hypoxia was induced by killing mice by nitrogen asphyxiation 15 min before the start of irradiation. *In vitro* aerobic irradiation was carried out on freshly prepared tumour-cell suspensions.

The results of *in vitro* irradiation and *in vivo* irradiation

Fig. 1 Survival curves for the xenografted tumour cells treated with single doses of γ radiation. ×, Cells irradiated as a single-cell suspension in aerobic conditions *in vitro* and assayed using the *in vitro* colony assay. The other data are for tumours irradiated *in situ*: ○ ●, assay immediately after irradiation in air-breathing mice; △ ▲, assay 18 h after irradiation in air-breathing mice; □ ■, assay immediately after irradiation in nitrogen-asphyxiated mice. Open symbols indicate the results of the diffusion chamber assay; closed symbols the results of the *in vitro* assay.
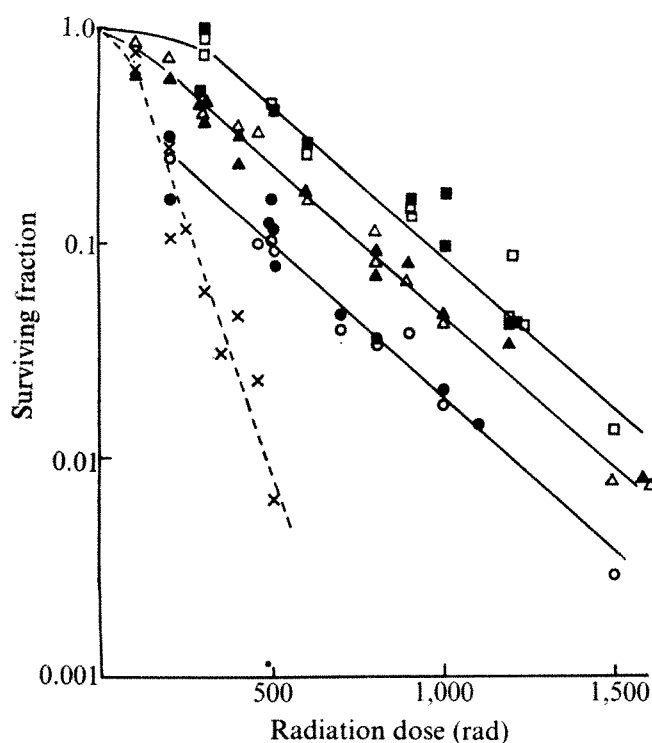
in air-breathing and hypoxic conditions are shown in Fig. 1. The two assays agree well and we are unable to detect any systematic differences between them. The *in vitro* cell survival curve has a D_0 of 96±9 rad. (D_0 is the dose that reduces survival by 63% on the exponential part of the curve.) The *in vivo* air-breathing survival curve for animals killed immediately after irradiation shows an initial sensitive component and a radioresistant component, the D_0 of which is 305±16 rad. The D_0 values for the terminal parts of the other two curves are not significantly different from this and we have therefore drawn the three curves parallel. From the vertical displacement of the hypoxic curve from the lowest air-breathing curve we calculate an hypoxic fraction of 0.25. The oxygen enhancement ratio is 3.1±0.4. The data obtained on tumours that were taken 18 h after irradiation, are displaced vertically from the other air-breathing curve by a factor of 2.2. This displacement shows that the fractional survival of tumour cells obtained at 18 h after irradiation was greater than when the cells were removed immediately. Two mechanisms may be involved: the repair of potentially lethal damage or the selective death or disappearance during the 18-h period of lethally damaged cells. Repair of potentially lethal damage has been reported in cell populations irradiated *in vitro* and *in vivo*[4,5]. The present data suggest that the 18-h curve is parallel to the other two curves down to radiation doses as low as 200 rad, but more data are required to confirm this.

Limited clinical experience with irradiation in pancreatic carcinoma suggests that this is not usually a radioresponsive tumour[6]. It is therefore of interest that the D_0 values which we have obtained are within the range of values determined using tumours of laboratory animals. The hypoxic fraction is also no less than has been found in experimental tumours, but the significance of this must be interpreted with caution because of the possible influence on this parameter of vascular stroma derived from the mouse. One interpretation of these data, however, is that the response of pancreatic carcinoma to radiotherapy is limited by the high proportion of hypoxic cells. This provides support for current efforts to improve clinical therapeutic response by the use of chemical radiosensitisers or high LET radiation.

V. D. COURTENAY
I. E. SMITH
M. J. PECKHAM
G. G. STEEL

Radiotherapy Research Department,
Divisions of Radiotherapy and Biophysics,
Institute of Cancer Research,
Sutton, Surrey, UK

1 Brown, J. M., and Berry, R. J., in *Symposium on the Effects of Radiation on Cellular Proliferation and Differentiation* (IAEA, Vienna, 1968).
2 Nias, A. H. W., in *The Biological and Clinical Basis of Radiosensitivity* (edit. by Friedman, M.), (Thomas, Springfield, Illinois, 1974).
3 Pickard, R. G. Cobb, L. M., and Steel, G. G., *Br. J. Cancer.* **31**, 36–45 (1975).
4 Little, J. B., Hahn, G. M., Frindel, E., and Tubiana, M., *Radiology*, **45**, 116–120 (1973).
5 Shipley, W. U., Stanley, J. A., Courtenay, V. D., and Field, S. B., *Cancer Res.*, **35**, 932–938 (1975).
6 Moertel, C. G., in *Cancer Medicine* (edit. by Holland, J. F., and Frei, E.), (Lea and Febiger, Philadelphia, 1973).

In vivo hyperthermia of Yoshida tumour induces entry of non-proliferating cells into cycle

HYPERTHERMIA (temperatures in excess of 40 °C) is at present receiving widespread interest as a potential method of treating cancer. High temperature can selectively destroy several types of cancer cell *in vitro* and *in vivo*[1-3], and extensive lysis has been achieved clinically in limb tumours maintained at 41.5–43.5 °C for 6–8 h (ref. 1). In spite of these encouraging findings, fundamental information is
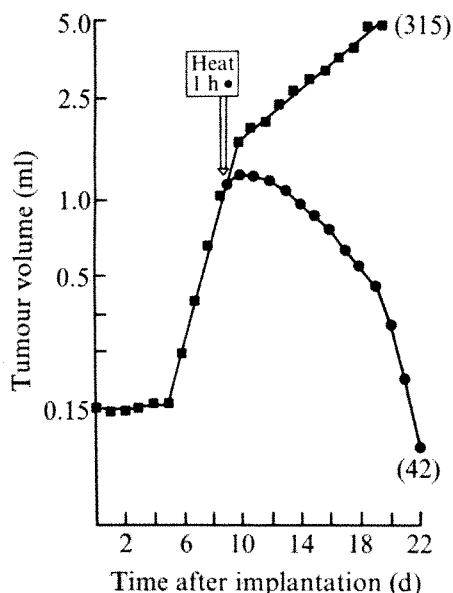
Fig. 1 Growth curve of the Yoshida sarcoma (■), and tumour volume changes after curative hyperthermia (intratumour temperature 42 °C for 1 h) on day 9 after implantation (●). Figures in brackets indicate the numbers of tumours (rats) used to construct the curves.

needed on the factors governing tumour response to heat before the place of this approach in cancer therapy can be defined. Inadequate heating of the Yoshida rat sarcoma (in terms of temperature and/or time relative to tumour volume) increased tumour metastasis and suggested an influence of population kinetics on the outcome of heating[4,5]. Work *in vitro* has shown hyperthermic cell killing to be preferentially cycle[6,7] and phase[8-11] oriented. Implication of cell kinetics in the response to heat is further suggested by work on the rabbit VX2 carcinoma. A single heat application (intratumour temperature 42 °C for 1 h) does not destroy this tumour, whereas three such treatments if applied within the tumour cell generation time are curative[12]. This report describes changes in the *in vivo* cell kinetics of the Yoshida sarcoma after curative heating at 42 °C, with emphasis on stimulation of non-proliferating tumour cells into cycle.

Details of the history and maintenance of the solid Yoshida sarcoma are described elsewhere[2]. For this work, the tumour was grown in the dorsum of one hind foot of the rats. After implantation of 0.1 ml homogenate the tumour grew progressively killing the host by metastasis in approximately 45 d. The tumour volume growth curve exhibited three distinct phases (Fig. 1)—an initial lag phase, followed by an exponential phase from day 5 to 10, and finally a slower growing 'tail' phase in keeping with the Gompertz growth function for solid tumours[13].

Tritiated thymidine (^{3}H-TdR; Radiochemical Centre, Amersham; 1 mCi ml[-1]; specific activity 21 Ci mmol[-1]) was diluted for injection with normal saline. ^{3}H-TdR was given intraperitoneally to the rats in the following doses: animals to be killed for the percentage labelled mitoses (PLM) and 1-h (flash) labelling work were given a single 400-μCi injection in 1 ml saline; animals for repeated tumour labelling experiments received 100 μCi ^{3}H-TdR every 6 h, and a further 100 μCi 1 h before killing. Autoradiographs of 4-μm paraffin-embedded hemisections of the tumours were prepared using the 'dipping' technique[14] with Ilford K-2 emulsion. After 14 d exposure time, the autoradiographs were processed with D-19 developer and F-5 fixer, and stained with haematoxylin and eosin. For the PLM curves, 500 anaphases and metaphases were counted 'blind', and for the flash and repeated labelling indices (% labelled cells) a

total of 2,000 cells were scored. Vincristine (Oncovin, Eli Lilly, Basingstoke, UK) was injected at a dose of 1 mg kg[-1] in 1 ml saline. Paraffin sections (4 μm) were stained as above and the mitotic index (% mitoses) determined by counting 4,000 cells.

The PLM curve of the 1.0–1.5 ml Yoshida sarcoma was constructed from results obtained with 60 tumours examined at 1-h intervals for 30 h. The median cell cycle times shown in Table 1 were obtained by graphical inspection of the curve at the 50% points. The flash-labelling index (I_L), or percentage labelled cells 1 h after exposure to ^{3}H-TdR, was 52.5%. These data were used to determine the growth fraction (fraction of the tumour cell population that is actively proliferating, I_P; ref. 15), assuming exponential growth and age distribution of the cells. The birth rate (K_B), or rate of production of new cells, was estimated by mitotic blocking with vincristine (Fig. 2), birth rate being the slope of the line[16]. Tumour doubling time (derived from the slope of the growth curve, Fig. 1) and birth rate data were used to estimate the cell loss factor[17] (or ratio of the rate of cell loss to birth rate).

Tumours of 1.0–1.5 ml were heated at 42 °C (intratumour temperature) for 1 h by immersing the foot bearing the tumour in a water bath. In these conditions, all tumours regressed completely within 14 d (Fig. 1) with cure of the host rats. The effect of hyperthermia on I_L is shown in Fig. 3, which provides an indication of the fraction of cells synthesising DNA over the 9 d after heat treatment. For the first 24 h after heat treatment, I_L was depressed but recovered to near control levels at 2 d; it then fell progressively to zero by 9 d as the tumour regressed. Changes occurring in the first 48 h after heat treatment were examined in more detail by repeated ^{3}H-TdR-labelling (Fig. 4). In controls repeatedly labelled every 6 h, the labelling index increased to a plateau level of almost 80% labelled cells after 8 h. With curative hyperthermia, a complex sequence of events ensued. Labelling was grossly depressed immediately after heating, but increased rapidly to a maximum of 60% after 8 h; an equally rapid decline in labelling index to a minimum of 5% at 14 h then preceded a final perturbed recovery to a plateau of approximately 60% labelled cells 36 h after heat.

The decline in labelling to 5% at 14 h after heat can only be due to cell death. Thus, more than 90% of the cells proliferating (P cells[15]) after hyperthermia had been lost by 14 h, the median cell cycle time of the tumour. This would imply that cells damaged at the time of heating had progressed through one cell cycle and died, possibly in mitosis. It seems unlikely that the rapid recovery of labelling and

Fig. 2 Mitotic accumulation after vincristine arrest in untreated Yoshida sarcoma. Cell birth rate (from slope) = 0.094/2 = 0.047 cells h[-1].
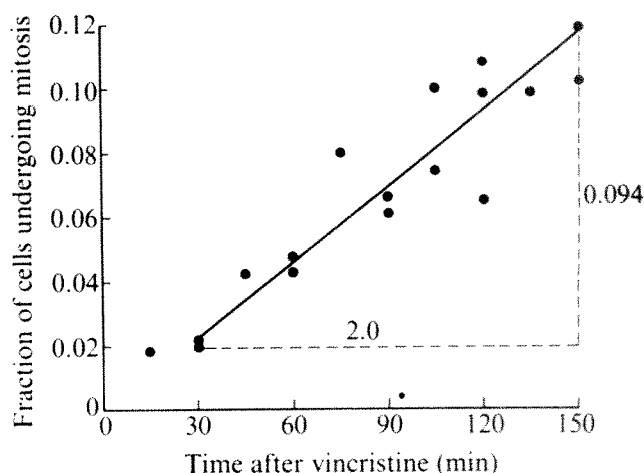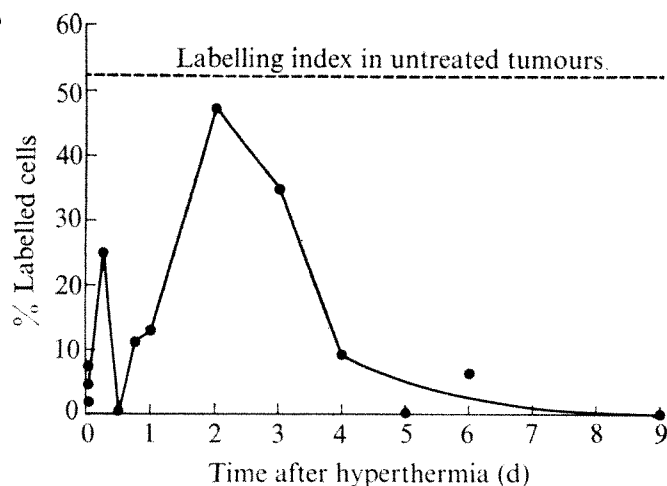
Fig. 3 ³H-TdR flash-labelling index in the Yoshida sarcoma after curative hyperthermia. From 30 min onwards, each point is the mean of two separate experiments.

Table 1 Cell kinetic parameters of the Yoshida sarcoma

Tumour volume (ml)				1.0–1.5
Doubling time (h)				36
Duration (h) of cell cycle and components:				

T_C	T_S	$T_{G_1}*$	T_{G_2}	T_M
(14.1)	(9.7)	(—)	(4)	(0.4)

Flash (1 h) labelling index	(I_L)	0.525
Growth fraction	(I_P)	0.678
Birth rate (cells h⁻¹)	(K_B)	0.047
Cell loss factor	(φ)	0.590

*Tumour has no detectable G_1 phase.

mitosis (as determined by vincristine blockade) from 14 to 36 h is caused by cells in cycle at the time of heat treatment. A more likely explanation is the entry into S phase of a population of cells not proliferating (Q cells) at the time of heat treatment; this would repopulate the tumour from 14 h onwards. Further evidence in favour of this is shown in Fig. 4. Repeated labelling was terminated at 14 h after heat, the time of maximum cell death; only a small increase in labelling was subsequently observed. When repeated labelling was commenced at 14 h, the labelling index recovered to the 60% region by 30 h after heat. These two experiments support the hypothesis that the recovery in proliferation from 14 h after heat in this tumour is mainly due to cells in the non-proliferating compartment at the time of heat treatment.

The PLM curve of the tumour 30 h after heat treatment provided durations for the cell cycle components that were not significantly different from those of the untreated tumour (Table 1). This is further substantiation that the changes in cell kinetics produced by curative hyperthermia are primarily in population size and the relationship of the P and Q cell compartments, rather than in cell generation times.

The data imply a preferential destruction of P cells in the Yoshida sarcoma by hyperthermia followed by the

recruitment of Q cells into cycle. Similar effects have been observed after radiation[18] or chemotherapy[19], and the timing of recovery has been used to plan fractionated therapy regimens. In spite of apparent kinetic recovery after heat treatment, the tumour regressed within 14 d with complete cure of the animals. The sensitivity of the Yoshida sarcoma to hyperthermia at volumes of 1.0 to 1.5 ml may be due to its high growth fraction ($I_P=0.678$) and the dominant position of the relatively sensitive S phase[10,11] in its cell cycle (9.7 out of 14.1 h). The failure of cells surviving the first 48 h after treatment to maintain the growth of the tumour implies an inability to repair sublethal damage and/or the operation of host factors in the destruction of the heat-treated tumour. ³H-TdR-labelled cells and mitoses were observed in tumour sections up to 6 d after heat treatment. This confirms *in vitro* findings[20] that several mitoses (up to 10 in the case of the Yoshida tumour) may occur before the expression of lethal hyperthermic damage.

J. A. DICKSON
S. K. CALDERWOOD

Cancer Research Unit,
University Department of Clinical Biochemistry,
Royal Victoria Infirmary,
Newcastle-upon-Tyne, UK

1 Cavaliere, R., *et al., Cancer*, **20**, 1351–1381 (1967).
2 Dickson, J. A., and Suzangar, M., *Cancer Res.*, **34**, 1263–1274 (1974).
3 Dickson, J. A., and Suzangar, M., in *Organ Culture in Biomedical Research* (edit. by Balls, M., and Monnickendam, M.), 417–461 (Cambridge University Press, Cambridge, London; New York and Melbourne, 1976).
4 Dickson, J. A., and Ellis, H. A., *Cancer Res.*, **36**, 1188–1195 (1976).
5 Dickson, J. A., and Ellis, H. A., *Nature*, **248**, 354–358 (1974).
6 Dickson, J. A., and Shah, D. M., *Eur. J. Cancer*, **8**, 561–571 (1972).
7 Kase, K., and Hahn, G. M., *Nature*, **255**, 228–230 (1975).
8 Rao, P. N., and Engelberg, J., *Science*, **148**, 1092–1094 (1965).
9 Sisken, J. E., Morasca, L., and Kibby, S., *Expl Cell. Res.*, **39**, 103–116 (1965).
10 Westra, A., and Dewey, W. C., *Int. J. radiat. Biol.*, **19**, 467–477 (1971).
11 Palzer, R. J., and Heidelberger, C., *Cancer Res.*, **33**, 422–427 (1973).
12 Muckle, D. S., and Dickson, J. A., *Br. J. Cancer*, **25**, 771–778 (1971).
13 Laird, A. K., *Natn. Cancer Inst. Monogr.*, **30**, 15–28 (1969).
14 Rogers, A. W., *Techniques of Autoradiography*, 253–273 (Elsevier, Amsterdam, London and New York, 1967).
15 Lala, P. K., *Meth. Cancer Res.*, **6**, 3–95 (1972).
16 Wright, N., Morley, A., and Appleton, D., *Cell Tissue Kinet.*, **5**, 351–364 (1972).
17 Steel, G. G., *Cell Tissue Kinet.*, **1**, 193–207 (1968).
18 Hermens, A. F., and Barendsen, G. W., *Eur. J. Cancer*, **5**, 173–189 (1969).
19 Wilcox, W. S., Griswald, D. P., Laster, W. R., Schabel, F. M., Jr, and Skipper, H. E., *Cancer Chemother. Rep.*, **47**, 27–39 (1965).
20 Palzer, R. J., and Heidelberger, C., *Cancer Res.*, **33**, 415–421 (1973).

Fig. 4 Repeated ³H-TdR labelling in untreated Yoshida tumours (△), tumours labelled 0–48 h after hyperthermia (●), tumours in which labelling was terminated at 14 h after heating (○), or tumours in which labelling was commenced at 14 h and continued to 48 h after heating (□).

Antibodies to melanoma cell and BCG antigens in sera from tumour-free individuals and from melanoma patients

THE *Mycobacterium bovis* strain BCG shares antigenic components with several neoplasms, including human malignant melanoma[1–5]. Other species of bacteria also have antigenic similarities with a guinea pig hepatoma[6]. Antibodies to mycobacteria and other bacterial antigens are universally present in sera from normal humans[7,8]. Since there is a broad range of shared or cross-reactive antigens among mycobacteria and unrelated microorganisms[9], it is

conceivable that an immune response by normal humans to mycobacteria has been induced by exposure to taxonomically unrelated microorganisms ubiquitous in the environment. Cytotoxic reactions by lymphoid cells from normal animals and humans to many kinds of tumour cells have been reported[10-15]. Antibodies in sera from patients with various neoplasms have often been observed that react with antigens derived from their own tumours or from those of the same histological type[16-18]. Humoral responses to tumour-associated antigens have also been noted in sera from normal humans, gibbons and mice[19-21]. We have investigated whether sera from normal, tumour-free humans contain antibodies which can react with melanoma-associated as well as with BCG antigens. If this were the case, perhaps immunological reactions by normal humans to certain tumours are widespread and reflect previous sensitisation by microorganisms.

BCG (Glaxo) was obtained from the US–Japan Cooperative Medical Science Program. Subcultures of BCG were grown, killed by heat, washed and disrupted by sonication. After centrifugation the resulting supernatant was separated and is referred to as BCG-SS (ref. 3).

Melanoma cells were grown in tissue culture and were provided by Dr G. Moore, Denver General Hospital[4]. They originated from a biopsy of a metastatic inguinal node of a male, aged 43 yr, whose primary lesion was on the heel of his right foot. They were treated with 3 M KCl as described before[3,4]. These KCl extracts have been used previously to investigate antigenic relationships between human malignant melanoma cells and BCG. It has been suggested that the antigenic determinants from this melanoma cell line which are shared with BCG are the group-specific cytoplasmic tumour antigens in melanoma cells[5].

The melanoma cell KCl extract and BCG-SS were labelled with ^{125}I using a modified chloramine T method as previously described[3,4] and are referred to as ^{125}I-mel and ^{125}I-BCG, respectively. The labelled antigens were diluted in 1 : 100 normal human serum in borate buffer so that their concentrations ranged from 0.01 to 0.008 μg N per 0.1 ml and so that the c.p.m. of 0.1 ml of labelled antigens used in the test procedures were 10,000–15,000.

The primary interactions between labelled test antigens and antibodies were analysed by precipitation of ^{125}I-labelled antigen–antibody complexes with anti-human IgG (anti-HGG). The amount of anti-HGG required to precipitate all

the HGG in a given test sample was determined as previously described[22]. Inhibition and absorption experiments were also carried out to evaluate the specificity of the observed reactions. Details of these procedures have been described before[3,4].

Sera from 63 patients with malignant melanoma were obtained shortly after surgery and before the patients were started on a chemoimmunotherapy programme. Thirteen patients had stage 1 disease, four had stage 2, 29 had stage 3 and 17 were stage 4. Ages ranged from 21 to 72 yr. There were 41 male and 22 female patients. Control sera were from 50 healthy subjects, most of whom were personnel at the National Jewish Hospital and Research Center (NJHRC). Sera were not obtained from subjects who had been in contact with human tumours in the laboratory. Ages ranged from 25 to 50 yr and there were 20 males and 30 females. Sera were obtained from two patients at the NJHRC who were under treatment for advanced, active tuberculosis.

When sera from all the patients and normal subjects were tested for their capacity to bind ^{125}I-BCG and ^{125}I-mel there was a wide range in their capacity to bind both radiolabelled antigens (Fig. 1a and b). As a group, sera from patients with melanoma bound ^{125}I-BCG to about the same extent as did sera from normal subjects. The capacity of sera from 17 patients with stage 4 disease to bind ^{125}I-BCG was somewhat less than that of normal sera. These differences, however, were not significant $(P<0.2)$. When the same sera were reacted with ^{125}I-mel, more antigen was bound by normal sera than by sera from the melanoma patients $(P<0.025)$. The overlapping of data between these two groups made it impossible to correlate disease and the amount of antigen bound by individual serum samples. When sera from patients with stage 4 were compared with normals they were found to have significantly lower $(P<0.005)$ binding values. Lower antibody levels to antigens have been shown when circulating antigen, soluble antigen–antibody complexes or anti-idiotype antibodies are present, and this may be the case in some patients with far advanced melanoma[18,23-26].

A series of experiments was then undertaken to evaluate the specificity of the binding data. The capacity of unlabelled BCG-SS to inhibit binding of ^{125}I-BCG and ^{125}I-mel by selected sera was examined. Sera from 11 melanoma patients and from seven normal subjects were preincubated

Fig. 1 Results of binding tests using ^{125}I-BCG (*a*) and ^{125}I-mel (*b*) expressed as percentage labelled antigens bound by 0.1 ml of a 1:5 dilution of whole serum. 1, All melanoma (63); 2, Stage 4 melanoma (17); 3, no disease (50).
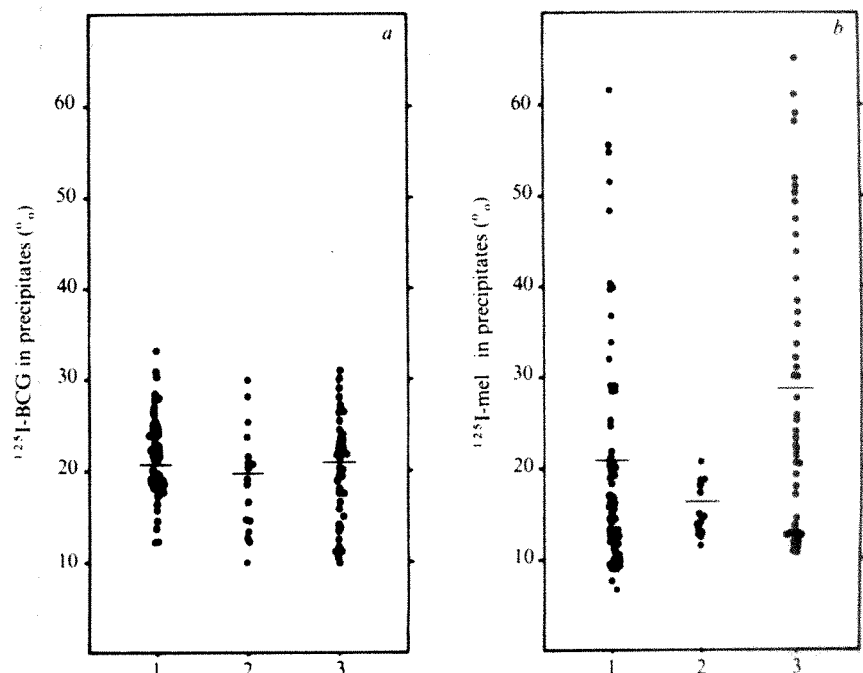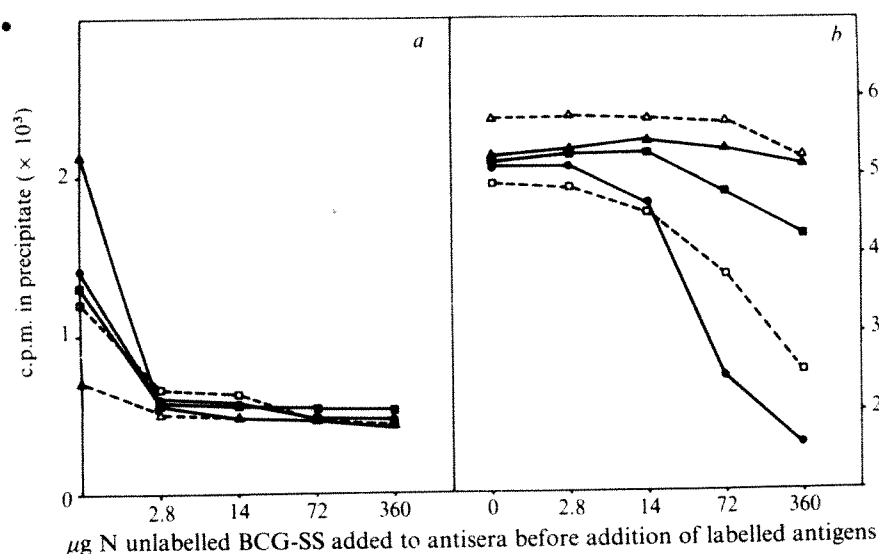
Fig. 2 Inhibition of binding of ^{125}I-BCG (*a*) and ^{125}I-mel (*b*) by sera from normal (open symbols) and melanoma (closed symbols) patients after preincubation of sera with dilutions of BCG-SS.

with 300 μg N unlabelled BCG-SS. Control samples were incubated with borate buffer. Twenty-four hours later, ^{125}I-BCG and ^{125}I-mel were added and binding capacities of the antisera were determined. The percentage decrease of binding is shown in Table 1. Unlabelled BCG-SS considerably inhibited the binding as expected to ^{125}I-BCG. To a lesser extent it also reduced the binding by most of these antisera to ^{125}I-mel. This is compatible with observations that some antigenic components are shared between BCG and melanoma cells[4].

Inhibition studies were carried out in greater detail using sera from three melanoma patients and two normal individuals. Serial dilutions of 0.1 ml of unlabelled BCG-SS were tested for their capacity to inhibit binding of 0.1 ml of ^{125}I-BCG and ^{125}I-mel to 1 : 5 dilutions of antisera. Controls consisted of 0.1 ml of borate buffer added to the diluted antisera. After overnight incubation, labelled BCG and melanoma antigens were added and binding to these antigens was determined. As little as 2.8 μg N unlabelled BCG-SS inhibited the binding by both normal and melanoma antisera to ^{125}I-BCG, and with increasing amounts there was slightly more inhibition (see Fig. 2*a* and *b*). When ^{125}I-mel

was the antigen there was similar inhibition of binding by four of the five sera tested, but more BCG-SS was required. As Fig. 2*b* shows, 72 μg N caused slight inhibition and 300 μg N were considerably more effective.

Sera from two melanoma patients, two normal subjects and two patients with active tuberculosis were then absorbed with 3×10^7 intact melanoma cells, or with 2×10^8 normal human spleen cells, and then reacted with ^{125}I-mel (Table 2). Sera from patients with tuberculosis were included because they generally have greater capacities to bind ^{125}I-BCG than do sera from normal subjects[7]. After absorption with melanoma cells there was a marked reduction in binding to ^{125}I-mel by all the sera. Absorption of the same sera with normal human spleen cells showed only slight decreases, except for normal serum No. 1, where there was a more moderate decrease in binding. Not shown in Table 1, there was no reduction in binding to ^{125}I-BCG after absorption with 3×10^7 melanoma cells. Unfortunately, greater numbers of melanoma cells were not available for additional absorptions.

Our previous studies and those of others have indicated that antibodies to mycobacteria and various unrelated bacteria and fungi are present in sera from humans with no history of exposure to these microorganisms[7,8]. In this study antibodies to a melanoma-associated antigen were detected in sera from the normal subjects tested. The binding and inhibition test results indicated that the reactions by normal sera were true immunological manifestations even though the specific stimulating antigen(s) was not identified. Because

Table 1 Percentage decreases of binding of ^{125}I-BCG and ^{125}I-mel by sera from normal subjects and patients with malignant melanoma after preincubation of sera with 300 μg N unlabelled BCG-SS

Antisera from	^{125}I-BCG antigen	^{125}I-mel antigen
Normal humans (No.)		
1	54.5	2.4
2	31.5	7.0
3	70.6	22.8
4	58.2	+3.2
5	29.3	0.4
6	55.5	42.7
7	36.6	19.9
Melanoma patients		
1	29.2	2.9
2	57.2	10.6
3	47.7	0.4
4	40.6	1.2
5	62.7	28.0
6	57.1	23.2
7	48.7	20.4
8	49.1	6.3
9	50.1	32.5
10	50.7	5.9
11	30.7	+1.6

Percentage decrease =
$$\left(\frac{\text{c.p.m. in precipitates after preincubation with BCG-SS}}{\text{c.p.m. in precipitates after preincubation with buffer}} - 1 \right) \times 100$$

Table 2 Binding of ^{125}I-mel by sera from normal persons, patients with malignant melanoma and tuberculosis after absorption with melanoma and spleen cells

Antisera from	c.p.m. in precipitates		
		Absorbed with	
	Unabsorbed	Melanoma cells*	Normal spleen cells†
Normal human (No.)			
1	7,546	1,827	5,681
2	7,983	1,847	7,742
Melanoma patient			
1	7,408	1,963	7,352
2	7,034	1,762	6,831
Tuberculosis patient			
1	5,373	1,695	● 5,269
2	1,834	1,520	1,798

*3×10^7 melanoma cells were used to absorb 1.0 ml of serum.
†2×10^8 normal human spleen cells were used to absorb 1.0 ml of serum.

of the antigenic relationships between BCG and melanoma cells[2,4,5], antibodies to melanoma-associated antigens might have been induced by exposure to mycobacteria or other microorganisms. Because the sharing of antigens between tumour cells and bacteria may be widespread[6], immune responses, whether cellular or humoral, may exist in normal humans to various tumours. Some of the frequently reported and unexplained cellular and humoral reactions to tumour antigens by normal persons[13–15] may have been induced by exposure to microorganisms. Immune responses by normal subjects[19–21] or animals to tumour antigens are often referred to as "natural" or "nonspecific". Results of this and previous studies[2,4,5] imply rather that immune reactions by normal subjects to melanoma-associated antigens can be the result of direct and specific stimulation by antigens of mycobacteria.

It is possible then that just as normal humans respond immunologically to bacterial antigens, they may also respond immunologically to some tumour antigens. If this is the case then immune responses by patients to tumour antigens may sometimes represent an elevation of an "isoimmune" state, rather than a new actively acquired immune response and may influence the course of neoplastic disease.

J.U.G. holds a career development award from the USPHS. This work was supported by grants from the NCI.

P. MINDEN
C. JARRETT
J. K. McCLATCHY

National Jewish Hospital and Research Center,
Denver, Colorado 80206

J. U. GUTTERMAN
E. M. HERSH

MD Anderson Hospital and Tumor Institute,
Houston, Texas 77025

1 Borsos, T., and Rapp, H. J., *J. natn. Cancer Inst.*, 51, 1085–1086 (1973).
2 Bucana, C., and Hanna, M. G., Jr, *J. natn. Cancer Inst.*, 53, 1313–1323 (1974).
3 Minden, P., McClatchy, J. K., Wainberg, M., and Weiss, D. W., *J. natn. Cancer Inst.*, 53, 1325–1331 (1974).
4 Minden, P., Sharpton, T. R., and McClatchy, J. K., *J. Immun.*, 116, 1407–1414 (1976).
5 Lewis, M. G., *et al.*, in *The Present Status of BCG in Cancer Immunotherapy* (edit. by Lamoureux, E. G., Portelance, V., and Turcotte, R.), (Grune and Stratton, New York, 1976).
6 Minden, P., Sharpton, T. R., and McClatchy, J. K., *Cancer Res.*, 36, 1680–1685 (1976).
7 Bardana, E. J., McClatchy, J. K., Farr, R. S., and Minden, P., *Clin. exp. Immun.*, 13, 65–77 (1973).
8 Minden, P., and Farr, R. S., *Chest*, 68, 749–750 (1975).
9 Minden, P., McClatchy, J. K., Cooper, R., Bardana, E. J., and Farr, R. S., *Science*, 176, 57–58 (1972).
10 Blair, P. B., Lane, M. A., and Yagi, M. J., *J. Immun.*, 112, 693–705 (1974).
11 Herberman, R. B., Nunn, M. E., Holden, H. T., and Lavrin, D. H., *Int. J. Cancer*, 16, 216–229 (1975).
12 Kiesling, R., Klein, E., and Wigzell, H., *Eur. J. Immun.*, 5, 112–121 (1975).
13 Tagasugi, M., Mickey, M. R., and Terasaki, P. I., *Cancer Res.*, 33, 2898–2902 (1973).
14 de Vries, J. E., Meyering, M., van Dongen, A., and Rumke, P., *Int. J. Cancer*, 15, 391–400 (1975).
15 Canevari, S., Fossati, G., and Della Porta, G., *J. natn. Cancer Inst.*, 56, 705–709 (1976).
16 Hellström, I., Hellström, K. E., Pierce, G. E., and Yang, J. P. S., *Nature*, 220, 1352–1354 (1968).
17 Lewis, M. G., *Med. Clinics of North America*, 56, 481–499 (1972).
18 Bodurtha, A. J., Chee, D. O., Laucius, J. F., Mastrangelo, M. J., and Prehn, R. T., *Cancer Res.*, 35, 189–193 (1975).
19 Bias, W. B., Santos, G. W., Burke, P. J., Mullins, G. M., and Humphrey, R. L., *Science*, 178, 304–306 (1972).
20 Aoki, T., *et al.*, *Science*, 19, 1180–1183 (1976).
21 Hanna, M. G., Jr, Ihle, J. N., Batzing, B. L., Tennant, R. W., and Schenley, C. K., *Cancer Res.*, 35, 164–171 (1975).
22 Minden, P., Farr, R. S., and Trembath, J., *Immunochemistry*, 12, 477–479 (1975).
23 Dixon, F. J., *The Harvey Lectures*, 58, 21 (Academic, New York, 1963).
24 Lewis, M. G., Phillips, T. M., Cook, K. B., and Blake, J., *Nature*, 232, 52–53 (1971).
25 Rowley, D. A., Fitch, F. W., Stuart, F. P., Kohler, H., and Cosenza, H., *Science*, 181, 1133–1141 (1973).
26 Lewis, M. G., Hartmann, D., and Jerry, L. M., *Proc. N.Y. Acad. Sci.*, (in the press).

Inhibition of tumour growth *in vivo* and *in vitro* by prostaglandin E

PROSTAGLANDIN E (PGE), both endogenous and exogenous, has been shown to inhibit significantly the rates of tumour-cell proliferation *in vitro*[1–4]. Suppression of PG biosynthesis with indomethacin[5,6] resulted in stimulation of cell replication, an effect which was readily reversed by the addition to the medium of small amounts of exogenous PGE_1 (10 ng ml^{-1}) (ref. 5). Corticosteroids have recently been shown to inhibit the biosynthesis of PGE by rheumatoid synovia[7] and mouse fibrosarcoma $HSDM_1$ (ref. 8) *in vitro*; the mechanism of this action seems to be by interfering with the release of arachidonic acid from phospholipids[9]. To evaluate further the possible role of PGE in the control of tumour-cell proliferation, we have studied the effects of hydrocortisone and indomethacin on the growth rate of B-16 mouse melanoma *in vitro*. The studies demonstrate that both compounds cause significant stimulation of tumour-cell replication *in vitro*, associated with inhibition of PG biosynthesis, albeit by two different mechanisms. In addition, we have extended the observations on the effects of exogenous PGE_1, by demonstrating that subcutaneous administration of 16,16-dimethyl-PGE_2-methyl ester significantly inhibits tumour growth *in vivo*.

The *in vitro* studies were all performed using B-16 melanoma cultured in modified McCoy's 5A medium, containing 15% foetal calf serum, glutamine (220 µg ml^{-1}), penicillin (100 U ml^{-1}) and streptomycin (0.1 mg ml^{-1}). Cultures were maintained in 75-cm^2 plastic T-flasks at 37 ± 0.5 °C in humidified 95% air, 5% CO_2, and media were changed on alternate days. Media containing PGE_1, indomethacin and hydrocortisone were sterilised by Millipore filtration. Control media contained identical concentrations of ethanol (< 0.01%). For counting, cells were removed using 0.1% trypsin, EDTA, 1.5 mg ml^{-1}, and mechanical scraping and were resuspended in medium. At least triplicate cell counts were performed using a haemocytometer; cell viability, as determined by vital dye exclusion[10] (Nigrosin 0.2%), ranged from 92.5 to 100% and was not influenced by the addition of exogenous PGE_1, indomethacin, or hydrocortisone. PG concentrations in media and supernatants were measured by radioimmunoassay after organic solvent extraction and silicic acid chromatography as described previously[5,11].

In an initial experiment, 2×10^5 B-16 cells were plated in duplicate flasks containing PGE_1 (1 µg ml^{-1}), indomethacin (10^{-8} M), or hydrocortisone (10^{-6} M). Exogenous PGE_1 caused a mean 24.4 ± 6.1% inhibition of cell replication (Fig. 1); inhibition was statistically significant at days 4 and 6 ($P < 0.05$ and $P < 0.001$, respectively). As in previous experiments, indomethacin caused a significant (35.5%) inhibition of PG concentrations in the media and resulted in a mean 17.0 ± 3.2% stimulation of cell replication, which was statistically significant at days 4 ($+18\%$; $P < 0.05\%$) and 6 (22%; $P < 0.005\%$). Hydrocortisone produced similar results, significantly stimulating cell proliferation and decreasing endogenous PGE concentrations in the media.

The relationship between the concentration of exogenous PGE and the degree of inhibition of cell proliferation was tested in duplicate experiments in which 0.5, 1.0, 2.0, 5.0 and 10.0×10^5 B-16 cells were exposed to PGE_1 (1 µg ml^{-1}) for 48 h. There was a gradual dose–response curve with higher ratios of PGE/cell causing more inhibition (31.4%, 37.5%, 25.8%, 22.8% and 22.4%, respectively).

The effect of hydrocortisone on the proliferation rate was investigated further by plating 2.5×10^5 cells from one stock culture in duplicate control media as well as duplicate media containing hydrocortisone at 0.3×10^{-5} M, 10^{-6} M, and 10^{-7} M. All three concentrations of hydrocortisone caused significant stimulation of cell replication (Table 1). Maximal stimulation (mean 36.5%) was achieved with 10^{-6} M hydrocortisone; comparable data for 0.3×10^{-5} M and 10^{-7} M were 18.1% and 32.2%, respectively. In spite of the large cell numbers achieved, there seemed to be no contact inhibition of growth. Total PGE biosynthesis (cells + media) was linear in control conditions and was inhibited significantly by hydrocortisone at all three concentrations (Fig. 2). Each of the three concentrations of hydrocortisone inhibited PGE synthesis approximately 50%. The effect of hydrocortisone on PGE concentrations was similar in media and cells, which rules out the possibility that hydro-

Table 1 The effect of hydrocortisone on B-16 proliferation *in vitro*

Hydrocortisone	Number of cells × 10⁶			
	0 h	48 h	96 h	144 h
0	0.25	0.40 ± 0.01	4.85 ± 0.55	9.95 ± 0.55
0.3 × 10⁻⁵ M	0.25	0.51 ± 0.03	5.80 ± 0.40	10.65 ± 0.75
10⁻⁶ M	0.25	0.67 ± 0.04	5.97 ± 0.02	11.78 ± 0.35
10⁻⁷ M	0.25	0.54 ± 0.06	6.55 ± 0.45	12.60 ± 0.90
10⁻⁶ M + PGE₁, 10 ng ml⁻¹	0.25	0.47 ± 0.04	4.65 ± 0.05	9.55 ± 0.65

The effect of hydrocortisone on B-16 proliferation *in vitro*. Data represent mean ± s.d. of duplicate determinations at each time point; at least triplicate counts were performed for each determination. Viability exceeded 95% in all cultures. Compared with control values, hydrocortisone caused mean stimulation of cell replication: 0.3×10^{-5} M, 18.1%; 10^{-6} M, 36.5%; 10^{-7} M, 32.5%; 10^{-6} M + 10 ng ml⁻¹ PGE₁, 3.2%.

cortisone only inhibited the release of PGE from cells. Using regression analysis, the degree of stimulation was not significantly related to the degree of inhibition of prostaglandin biosynthesis ($P > 0.05$). Therefore, in order to verify that the effect of hydrocortisone on stimulation of cell replication was, in fact, a result of its induced inhibition of prostaglandin biosynthesis, an identical experiment was carried out in which cells were grown in media containing both 10^{-6} M hydrocortisone and exogenous PGE₁, 10 ng ml⁻¹ (the approximate endogenous concentration in control conditions). Small concentrations of exogenous PGE₁ normalised the growth rate (Table 1).

Fig. 1 Effect of PGE₁, indomethacin and hydrocortisone on B-16 replication *in vitro*. Data represent means of duplicate flasks at each time point; for each determination at least triplicate counts were performed. ▲, Controls; ■, hydrocortisone 10^{-6} M; ○, indomethacin 10^{-8} M; ●, PGE₁, 1 μg ml⁻¹.
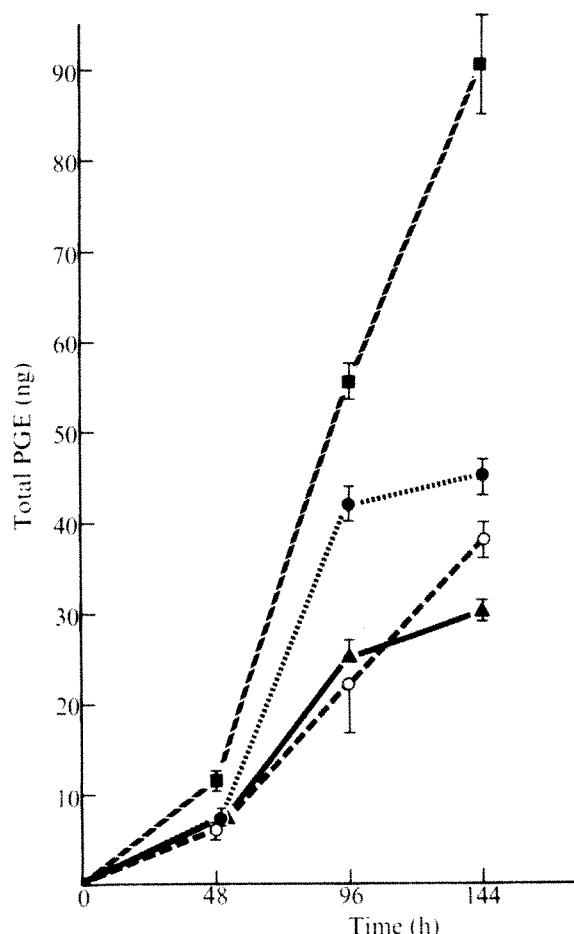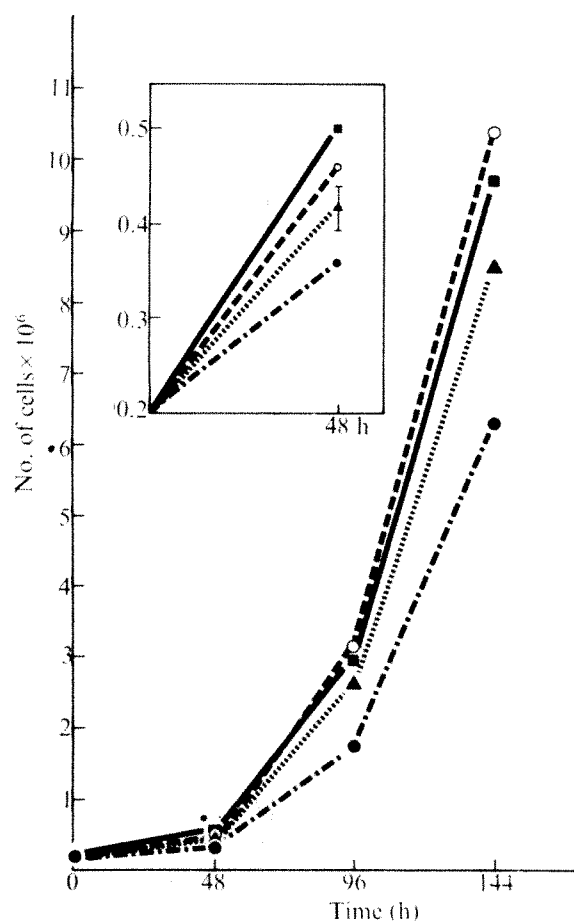
Fig. 2 Effect of hydrocortisone on PGE biosynthesis. Data represent total (media + intracellular) concentrations mean ± s.e.m. Each determination was based on duplicate radioimmunoassay determinations on replicate samples. Hydrocortisone: ■, 0 (control); ○, 0.3×10^{-5} M; ●, 10^{-6} M; ▲, 10^{-7} M. Compared with the controls, hydrocortisone produced significant ($P < 0.05$) inhibition of PGE biosynthesis, 63.0%, 45.1% and 57.7%, respectively. Hydrocortisone produced slightly less inhibition of PGF synthesis; the PGE/PGF rates averaged 3.12 in the control experiments and 2.39 in the hydrocortisone experiments.

In an experiment designed to evaluate if hydrocortisone interfered with the cyclo-oxygenase system, homogenates were prepared in which 5×10^5 sonicated cells were mixed with glutathione (55 μg ml⁻¹), hydroquinone (0.55 μg ml⁻¹), EDTA (5.5 μg ml⁻¹), and arachidonic acid (500 μg ml⁻¹) as previously described[12]. After 1 h of incubation at 37 °C in 95% air and 5% CO₂, the homogenates were extracted for measurement of PGE concentrations. In the control conditions, homogenates of 5×10^5 cells synthesised 9.7 ng of immunoreactive PGE. Inclusion of hydrocortisone at a final concentration of 10^{-6} M did not alter PGE synthesis (9.6 ng). On the other hand, indomethacin (10^{-8} M) produced 92.9% inhibition of the conversion of arachidonic acid to PGE₂ (0.74 ng).

In *in vivo* studies, 37 3-week old litter mate female C57BL/6J mice (Jackson Laboratories) were injected subcutaneously with 1×10^6 viable B-16 melanoma tumour cells suspended in 0.2 ml of McCoy's medium containing penicillin and streptomycin but no foetal calf serum. They were divided randomly into control ($n = 19$) and test ($n = 18$) groups and within 2 min received a subcutaneous injection in the same area of 0.1 ml of 0.9% sodium chloride containing either 25% absolute ethanol (control group) or an ethanolic solution containing 5 μg of 16,16-dimethyl-PGE₂-methyl ester (test group). The

Table 2 The effect of 16,16-dimethyl-PGE$_2$-methyl ester on B-16 growth *in vivo*

No.	Control group Tumour weight* (mg)	Cells per g tumour (×10⁶)	Body weight (g)	Total cells in tumour† (×10⁶)	No.	PGE group Tumour weight* (mg)	Cells per g tumour (×10⁶)	Body weight (g)	Total cells in tumour† (×10⁶)
1	—	—	23.0	—	1	—	—	23.5	—
2	230.1	—	20.0	59.46	2	—	—	23.0	—
3	230.8	284.0	19.0	59.64	3	30.6	—	27.0	5.63
4	1,008.7	266.6	20.0	260.65	4	37.0	—	25.5	6.81
5	685.7	—	18.0	177.00	5	—	—	9.0	—
6	834.9	236.8	11.0	215.74	6	183.0	—	15.0	33.70
7	428.0	—	12.2	110.59	7	63.1	—	19.7	11.62
8	477.3	230.8	11.3	123.33	8	481.6	206.3	18.5	88.71
9	487.1	—	14.9	125.86	9	455.0	160.7	15.9	83.81
10	125.9	—	14.5	32.53	10	45.6	—	16.9	8.39
11	465.7	—	13.8	120.33	11	151.7	185.8	11.9	27.94
12	508.0	—	12.0	131.26	12	24.3	—	15.5	4.47
13	492.3	—	14.0	127.21	13	—	—	20.1	—
14	243.7	—	13.0	62.97	14	30.1	—	21.3	5.54
15	432.5	—	11.0	111.75	15	—	—	15.4	—
16	818.5	—	14.8	211.50	16	64.1	—	16.7	11.80
17	276.9	—	17.8	71.55	17	179.7	—	18.8	93.10
18	988.6	273.8	14.5	255.45	18	231.1	—	17.9	42.57
19	136.3	—	19.2	35.22					
Mean (18)	492.8	258.4	15.4	127.3	(13)	152.0	184.2	18.08	28.0
±s.e.	±64.9	±10.4	±0.8	±16.7		±43.4	±13.1	±1.1	±8.0
Mean (19)	466.9	—	—	120.6	(18)	109.8	—	—	20.2
±s.e.	±66.6			±17.2		±35.1			±6.5

The effect of 16,16-dimethyl-PGE$_2$-methyl-ester on B-16 growth *in vivo*.
*Tumour weights represent wet weights immediately after collection.
†Total cells per tumour was determined by tumour weights × mean cell number per g of tumour in each group.

identical injection procedure was repeated daily. Tumours were apparent approximately 10 d after innoculation in both groups. After 18 d the mice were killed with ether and the tumours were collected and weighed. No necrosis was apparent in any tumour in either group. Tumour cells were separated by passage through a 120 mesh stainless steel grid. After washing twice with medium, the cells were resuspended and counted as described previously. As shown in Table 3, 18 of 19 (95%) mice in the control group developed subcutaneous tumours which weighed an average of 492.8 ± 64.9 mg. In contrast, only 13 of 18 (72%) of the mice in the test (16,16-dimethyl-PGE$_2$-methyl ester) group developed identifiable tumours. Using a χ^2 test, this difference was statistically significant with a P value of < 0.05. Furthermore, among the 13 tumour-bearing animals, the mean tumour weight was 152.0 ± 43.4 mg representing 69.2% inhibition; compared with the mean weight of the tumour-bearing controls, this difference was also statistically different with a P value of < 0.001. Based on the mean number of cells per g of tumour in each group, it was possible to plot the fate of the 10⁶ injected cells. In the control group, the tumours represented an average of 120.6 × 10⁶ ± 17.2 cells per mouse (including the tumour-free animal) whereas in the test group, they represented 20.2 × 10⁶ ± 6.5 cells per mouse. This difference (inhibition by 83.2%) was statistically significant ($P < 0.001$). At the completion of the experiment, the control mice weighed an average of 15.4 ± 0.8 g, and the test group 18.1 ± 1.1 g ($P < 0.05$); there was no correlation between tumour weight and the body weight of the mice.

These studies have further implicated PGE in the control of tumour cell growth. Inhibition of PGE biosynthesis, either by indomethacin or by hydrocortisone results in stimulation of cell proliferation; the effect of both of these compounds can readily be reversed by the addition of small (subthreshhold) amounts of PGE$_1$. The effect of PGE$_1$ in inhibiting proliferation is, in general, dose related. These *in vitro* observations are further substantiated by the *in vivo* experiments, in which subcutaneous (and after the appearance of tumours, intra-tumour) administration of a long-acting PG analogue resulted in significant suppression of tumour growth.

We acknowledge the technical assistance of D. Lazan, D.

M. G. SANTORO
G. W. PHILPOTT
B. M. JAFFE

Department of Surgery,
Washington University School of Medicine,
St Louis, Missouri 63110

1 Prasad, K. N., *Nature new Biol.*, 236, 49–52 (1972).
2 Johnson, G. S., Pastan, J., Perry, C. V., Otten, J., and Willingham, M., in *Prostaglandins in Cellular Biology* (edit. by Ramwell, P. W., and Pharris, B. B.), 195–205 (Plenum, New York, 1972).
3 Thomas, D. R., Philpott, G. W., and Jaffe, B. M., *J. surg. Res.*, 16, 463–465 (1974).
4 Hamprecht, B., Jaffe, B. M., and Philpott, G. W., *FEBS Lett.*, 36, 193–198 (1973).
5 Thomas, D. R., Philpott, G. W., and Jaffe, B. M., *Expl Cell Res.*, 84, 40–46 (1974).
6 Vane, J. R., *Nature new Biol.*, 231, 232–235 (1971).
7 Kantrowitz, F., Robinson, D. R., McGuire, M. B., and Levine, L., *Nature*, 258, 737–739 (1975).
8 Tashjian, A. H., Jr, Voelkel, E. F., McDonough, J., and Levine, L., *Nature*, 258, 739–741 (1975).
9 Hong, S. L., and Levine, L., *Proc. natn. Acad. Sci. U.S.A.* (in the press).
10 Phillips, H. J., and Terryberry, J. E., *Expl Cell Res.*, 13, 341–347 (1957).
11 Jaffe, B. M., Behrman, H. R., and Parker, C. W., *Clin. Invest.*, 52, 398–405 (1973).
12 Cohen, F., and Jaffe, B. M., *Biochem. biophys. Res. Commun.*, 55, 724–729 (1973).

Signal for cell fusion

THE mating of the gametes of *Ulva mutabilis*[1] occurs essentially as described for *Chlamydomonas*[2,3]. When the gametes are mixed, they cluster and agglutinate with the tips of their flagella. Pairs consisting of a (+) and (−)-gamete leave the cluster, the cell bodies are brought together and fusion initiated. In *Chlamydomonas* agglutination apparently creates a signal which triggers the later events of cell wall lysis, mating structure activation and cell fusion[4]. The gametes of *Ulva* lack a cell wall[5] and have apparently no elaborate mating structure[2,5,6]. This note shows that the signal in *Ulva* activates particular areas of the membrane for fusion only for a limited time, and if fusion does not occur then, the ability to fuse is lost permanently.

Figure 1a shows that fusion is finished within 3–4 min both at 10 °C and 22 °C, but that the numbers which fuse are considerably lower at 10 °C than at 22 °C. The data in Figure 1b indicate that the number is constant above 15 °C,

decreases rapidly with decreasing temperature around 10 °C and reaches zero at 0 °C.

The membranes of the two gametes fuse in a region below the flagella and the fused area increases in size as the two gametes merge into each other. The merging is temperature dependent (Fig. 2, insert), and an Arrhenius plot of the initial rate of merging reveals a low activation energy above 15 °C and a higher below 15 °C. The rate most probably depends on the physical state of the membrane, and it is therefore suggested that the membranes change properties around 15 °C. This is also the temperature below which the number of fusers decreases. The decrease can be explained if fusion can take place only for a limited period of time and if the probability for fusion is temperature dependent below 15 °C.

When the two sexes are incubated separately at low temperature for 30–40 min, and mixed after raising the temperature to 22 °C, no reduction in the number of fusers occurs. Consequently the reduction in the ability to fuse develops after the two sexes have been brought together.

Fig. 1 *a*, The percentage of zygotes formed as function of time at 22 °C (●) and 10 °C (+). The two mating types were mixed at time zero. *b*, The final percentage of zygotes as a function of temperature. The percentage at 22 °C is taken as 100. Gametogenesis was induced as described by Nordby and Hoxmark[12]. The experiments were performed at a cell density of 10^6–10^7 gametes ml^{-1}. At different times after mixing the (–) and (+)-gametes the process was stopped by adding an equal volume of 2.5% glutaraldehyde in 0.2 M cacodylate buffer pH = 8, and 100–400 gametes (*g*) and zygotes (*z*) were counted in the phase contrast microscope. The percentage of fusion is given as $200\ z/g + 2z$.
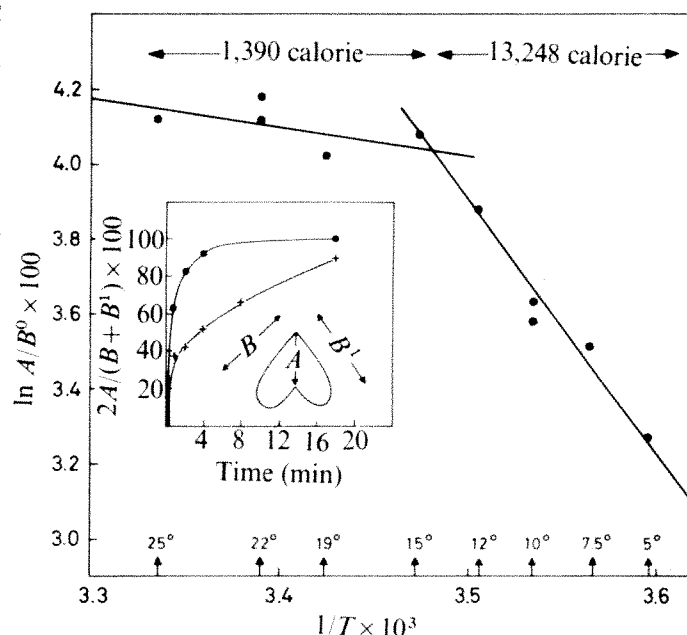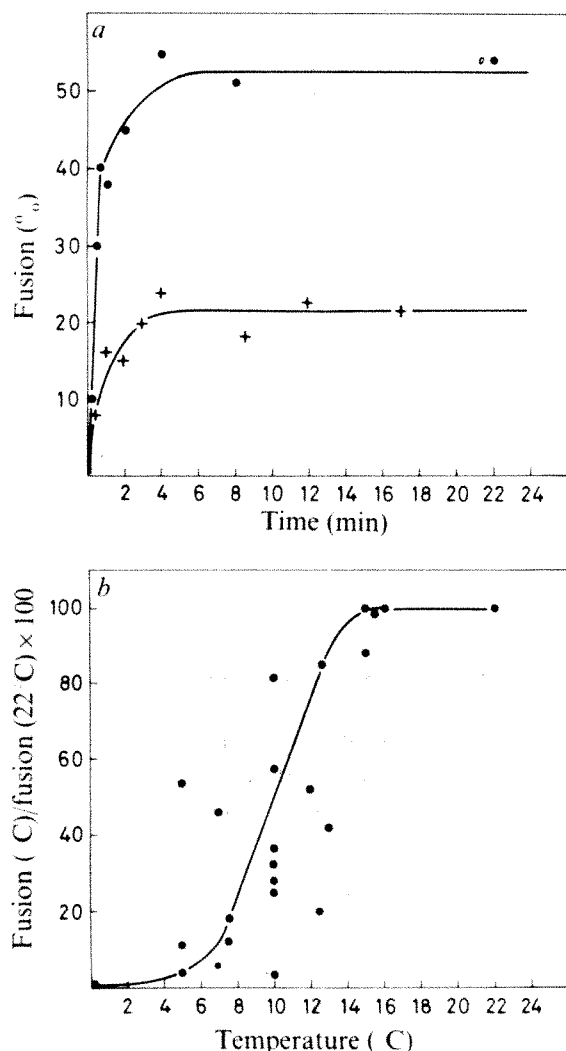
Fig. 2 Relationship between temperature and the rate of merging. The degree of merging is measured by the length of line A (see insert) and is given as the percentage of the average (*n* = 25) length of the fused gametes ($\frac{1}{2}(B+B^1)$ = B^0). The insert gives the percentages of merging at different times at 10 °C (+) and 22 °C (●). The percentages obtained 1 min after mixing are used as an approximate measure of the initial rate of merging, and such rates are given in the Arrhenius plot. The lines are fitted to the points by the method of the least squares. Above 15 °C the slope is presumably not significantly different from 0, which means zero activation energy. The zygotes were fixed for 4–5 h (see legend to Fig. 1), filtered on to a Nucleopore filter CPRO 1300 0.8 μm, dried by the critical point drying method[13] and coated with Au. The measurements were made directly on the screen of the Jeol ISM-S1 scanning electron microscope (× 10,000), operated at an angle of 0°.

Figure 3 indicates that exposure to 0 °C for only 3–5 min after mixing the gametes reduces the number of fusers to 50%, and exposure for 11 min almost totally prevents fusion.

The effect of cold is simulated at room temperature by cocaine, which is believed to be a membrane stabiliser and prevents virus-induced fusion in mammalian cell strains[7]. When gametes were treated with 5×10^{-3} M cocaine for 15 min, washed three times and mixed immediately, 62% fused. When the gametes were mixed after 5 min and washed after 10 min, the corresponding value was only 2%.

Cocaine at this concentration enables agglutination to occur. Observations also show that at 0 °C the two mating types agglutinate and stay in contact by the tips of their flagella for long periods. It is therefore suggested that this contact starts a process which within minutes makes the gametes incapable of fusion.

To investigate the role of flagellar contact on the fusion process, the flagella were removed with a 15-ml Dounce homogeniser (Pistil B) and the deflagellate gametes mated with biflagellate ones of the opposite sex. The degree of deflagellation could be regulated by the number of strokes, and it appeared that the number of zygotes formed was roughly proportional to the number of biflagellate gametes in the treated sample.

In fixed gametes the flagella stick out from the cell body and their presence is easy to observe. In the zygotes the flagella are directed backwards close to the cell body and it is not always easy to ascertain that all four are present. On inspecting 50 normal zygotes in the phase contrast microscope four flagella were observed in 64%, three flagella in 26%, two flagella in 10% and no zygotes carried one or no flagella. When deflagellated (–)-gametes (no flagella: 73%, one flagellum: 24%, two flagella: 3%) were

Table 1 The effect of con A on pairing

Mating type of the deflagellated gametes	Con A	Mating type of the flagellated gametes	Pairs consisting of: Gametes of the same mating type	Gametes of different mating type	Zygotes	Aggregates
+	+	—	3	70	11	Few
+	—	—	3	78	3	Few
—	+	+	59	0	4	Many
—	—	+	14	51	4	Few

(+) and (—)-gametes were deflagellated with six strokes of the Dounce homogeniser (80% gametes with no flagella, 15% with one flagellum, 5% with two flagella) and pretreated with fluorescent con A (100 µg ml⁻¹)[11], washed four times and then challenged with untreated biflagellate gametes of the opposite sex. The controls were done with untreated deflagellate gametes. Ten minutes after mixing the samples were concentrated in the centrifuge, suspended in 50 µl medium and fixed in 4% formalin. The samples were examined for 10 min each in the fluorescence microscope and the number of pairs and zygotes recorded.

mated with normal (+)-gametes, the corresponding numbers ($n=100$) were 7%, 66%, 21%, 3%, 3%. The high occurrence (66%) of zygotes with three flagella in the experiment compared with the control (26%) indicates that a uniflagellate gamete can participate in zygote formation. If the probability of not observing a flagellum which is present is the same in the experiment and the control, the data also suggest that a gamete without any flagella has a very low chance of fusing.

There seemed, however, to be a tendency for these gametes to pair with others without fusion, as indicated by the preponderance of pairs with two flagella in the experiment above. Four per cent of pairs carried four flagella, 22% carried three flagella, 62% carried two flagella, 12% carried one flagellum and 4% had no flagella. In the pairs the individuals adhere to each other just below the flagella where fusion is normally initiated. When both mating types are deflagellated and brought together with repeated centrifugations many pairs but no zygotes are observed.

In *Chlamydomonas*, concanavalin A (con A) causes isoagglutination of one or both sexes depending on the species[8-10]. In *Ulva*, con A (Sigma) isoagglutinates the flagellate gametes of both sexes at a concentration of 25 µg ml⁻¹ or above, but nevertheless it is possible to reveal a difference between the two kind of gametes. A 50% inhibition of zygote formation is obtained with 4 µg ml⁻¹ con A. The inhibition is due to con A-specific sites at the (—)-gametes, as shown by an experiment in which the gametes were pretreated for 5 min at a concentration of 25 µg ml⁻¹ con A, washed with medium four times, and mixed with untreated gametes of the opposite sex, which had not been treated with con A but had been washed four times. When the con

Fig. 3 The loss of fusion ability with time. The gametes were precooled at 0 °C for 25 min, mixed at zero time and at the times indicated 10 µl of the mixture was dropped on to a slide at room temperature and the percentage fusion recorded 15 min later.

A-treated gametes were mating type (+), the percentage of fusion was normal (experiment 53 control 45%). When mating type (—) was pretreated, only 1% fused.

If deflagellate gametes were treated with fluorescein-labelled con A, it could be demonstrated that pairing of (+) and (—)-gametes was prevented only by con A treatment of the deflagellate (—)-gametes (Table 1). While the flagellate gametes of both sexes are isoagglutinated in the presence of con A, treatment and subsequent washing causes isoagglutination of only the (—)-sex. These observations suggest that the region of the membrane which pairs and later fuses, represents a special part which is different in the two kinds of gametes.

In conclusion, the results can be explained if contact between the tips of the flagella creates a signal which activates a particular area of the membrane of each gamete for a limited period of time. If the two areas, which are different, are brought into contact, there is a given probability for fusion. This probability is dependent on the physical state of the membrane, and decreases with temperature below 15 °C or if membrane stabilisers like cocaine are added. If fusion does not occur during the defined period, the gametes separate and have lost the ability to fuse for good, either because the membranes are irreversibly deactivated or because a new signal cannot be generated.

ARNE LØVLIE
EVA BRYHNI

*Institute of Zoology,
University of Oslo,
Box 1050, Oslo 3, Norway*

1. Føyn, B., *Arch. Protistenk.*, 102, 473–80 (1958).
2. Cavalier-Smith, T., *Protoplasma*, 86, 1–18 (1975).
3. Wiese, L., in *Fertilization* II (edit. by Metz, C. B., and Monroy, A.), 135–188 (Academic, New York, 1969).
4. Goodenough, Ursula W., and Weiss, Richard L., *J. cell. Biol.*, 67, 623–637 (1975).
5. Bråten, T., *J. Cell Sci.*, 9, 621–635 (1971).
6. Friedmann, I., Colwin, A. L. and Colwin, L. H., *J. Cell Sci.*, 3, 115–128 (1968).
7. Poste, G. and Reeve, P., *Expl Cell Res.*, 72, 556–560 (1972).
8. McLean, R. J., and Brown, R. M., *Devl Biol.*, 36, 279–285 (1974).
9. Wiese, L. and Hayward, P. C., *Am. J. Bot.*, 59, 530–536 (1972).
10. Wiese, L., and Shoemaker, D. W., *Biol. Bull. Woods Hole*, 138, 88–95 (1970).
11. Wayne Smith, C., and Hollers, James C., *J. Reticuloendoth. Soc.*, 8, 458–464 (1970).
12. Nordby, O., and Hoxmark, R. Chr., *Expl Cell Res.*, 75, 321–328 (1970).
13. Anderson, T. F., *Trans. N.Y. Acad. Sci.*, 13, 130–134 (1951).
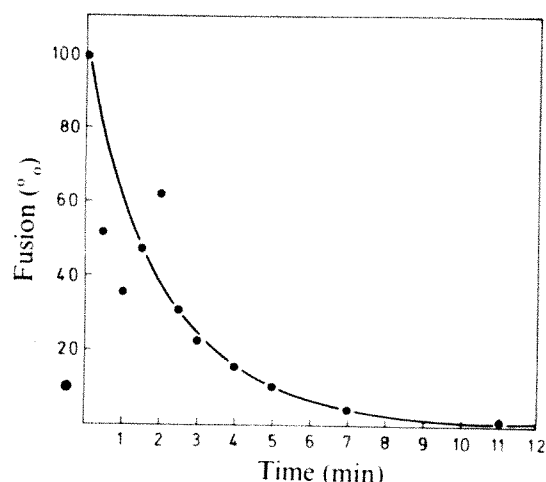
Recycling of dissolved plasma membrane components as an explanation of the capping phenomenon

IN a recent paper in *Nature*, Bretscher[1] proposed that the capping phenomenon[2,3] is caused by a continuous, directional flow of the lipid components of the plasma membrane across the cell surface from one end to the other, while the membrane proteins remain dispersed by random thermal motion and therefore do not normally participate. Bretscher's explanation of capping is that the cross linking of membrane proteins by attached antibodies and lectins causes them to be swept along by this lipid

current and collected at the "uroid" portion of the cell surface, where he proposes that the membrane lipid is taken into the cytoplasm (by way of a "molecular filter") for return to the opposite side of the cell as cytoplasmic vesicles. In this way Bretscher can account for the selective accumulation into caps of certain classes of surface antigen, without needing to postulate any communication between the surface and the interior. Cross-linked antigens would simply be swept into caps by the lipid current, while uncross-linked antigens would remain dispersed, just as is actually observed.

As an alternative to Bretscher's theory, I propose that capping results from a directional flow of the entire plasma membrane, including not only its lipid but also its integral protein and carbohydrate components. This alternative hypothesis is an extension of the theory previously proposed by Abercrombie et al.[4] to explain fibroblast locomotion and has been developed primarily on the basis of studies of cell locomotion and particle transport[5-7]. The basic postulate of my hypothesis is that all the components of the plasma membrane can dissolve into the cytoplasm at least to a slight extent and that in all motile tissue cells true thermodynamic equilibrium exists between the plasma membrane and its components dissolved in the cytoplasm.

According to LeChatelier's principle of mobile equilibria, forces imposed on the reactants of an equilibrium will cause the equilibrium to shift in whichever direction tends to absorb the imposed force[8]. I propose therefore that local stretching of the plasma membrane should shift this equilibrium in favour of the insertion (assembly) of dissolved membrane components into the plasma membrane. Consequently membrane assembly should be concentrated where the plasma membrane is most stretched. Conversely, membrane disassembly, that is, the dissolving of components back into the cytoplasm, should be favoured wherever the membrane is least stretched. In this way, a steady longitudinal pull applied to the plasma membrane by the cytoplasmic microfilaments would result in reassembly of membrane along the leading margin of the cell, followed by rearward flow of this membrane, with disassembly occurring in the more posterior portion of the cell surface. The disassembled (dissolved) membrane components would then be free to diffuse forward through the cytoplasm back to the reassembly area.

Fig. 1 Diagram of the proposed pattern of plasma membrane movement during lymphocyte capping and tissue cell locomotion. Along the leading margin of the cell, plasma membrane components dissolved in the cytoplasm are rapidly intercalated into the surface (membrane reassembly). The plasma membrane flows rearward from this assembly area, exerting traction against any adhering objects and thereby propelling the cell forward. At the rear "uropod" area of the cell, membrane components re-dissolve in the cytoplasm (membrane disassembly) and diffuse through the cytoplasm back to the front. Bound antibodies, lectins, and so on, are carried to the uropod area by the rearward membrane flow, and being unable to dissolve in the cytoplasm themselves, accumulate there to form caps.
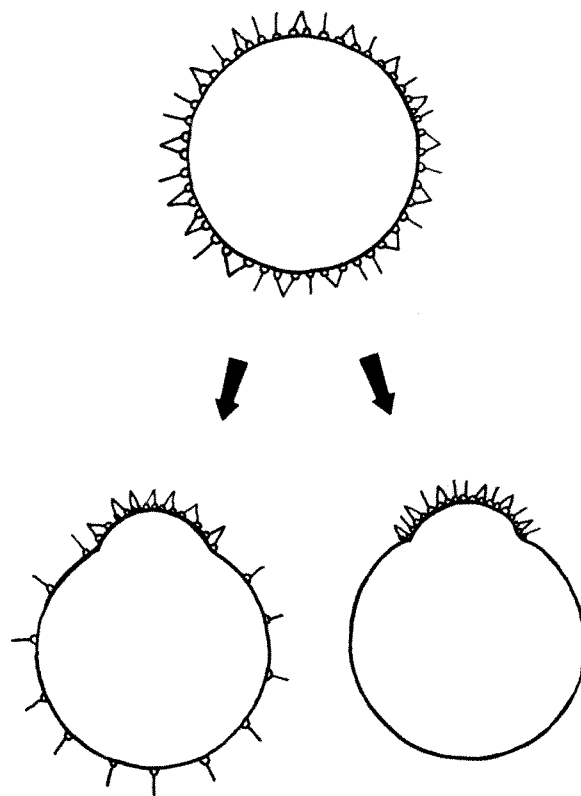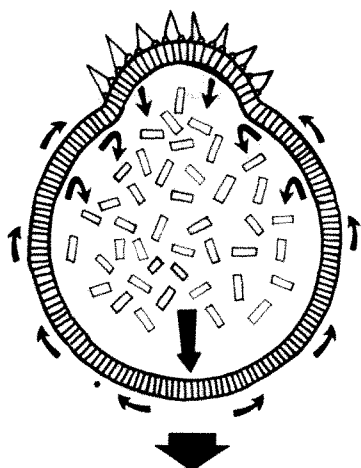
Fig. 2 Diagram of the experiment suggested to distinguish between the membrane recycling theory proposed in this paper and Bretscher's previously proposed theory of directed lipid flow. Lymphocytes are simultaneously exposed to multivalent antibodies against one membrane antigen and to univalent antibodies (Fab fragments) against another of their membrane antigens. The membrane recycling theory predicts that both multivalent and univalent antibodies should accumulate together in the same cap, while Bretscher's hypothesis leads to the expectation that the multivalent antibody should form a cap, while the univalent antibody should remain dispersed.

This directional flow of plasma membrane is envisaged as being responsible for locomotion as well as for capping. Any materials (antibodies, lectins, or small solid particles) which become attached to the surface of a tissue cell undergoing such membrane flow should be carried along to the uroid area of membrane disassembly. Being unable to dissolve in the cytoplasm, these attached materials should remain at the surface, accumulating there to form the familiar "cap". If the attached antibodies and lectins are bound strongly enough to certain components of the membrane, then these components should also be trapped at the surface and accumulated at the cap. Unbound membrane components would, of course, continue to be recycled without accumulating in the cap, though they would pass repeatedly through the disassembly area. In this way my hypothesis can account for the selective accumulation of bound antigens without the need to postulate a molecular filter and also explains why the caps are ordinarily located at the rear uropod area of the cell surface.

The postulate that all components of the plasma membrane, including the lipids, can dissolve into the cytoplasm may at first seem highly implausible or even unacceptable to some readers. Considerable biochemical evidence is, however, accumulating that phospholipids are rapidly interchanged between the several membrane systems of cells. For example, there is rapid equilibration of isotopically labelled phospholipids between mito-chondrial and microsomal membranes[9,10]. This interchange is rapid, reversible and temperature dependent, but does not require metabolic energy, indicating that it represents a true thermodynamic equilibrium. If phospholipids can disassemble from one class of cytoplasmic membrane and reassemble into

another, then it is less surprising that lipid and other membrane components can be disassembled from one part of the plasma membrane and reassembled back into another part of this same membrane, as required by the hypothesis proposed here.

One apparent contradiction to this hypothesis is the observation that univalent antibodies (Fab fragments) and univalent lectins can be bound to lymphocytes without accumulating into caps[2]. However, this observation does not actually contradict the hypothesis because the lymphocytes in question were non-motile and not even attached to a substratum and they were not shown to be simultaneously accumulating other attached materials into caps. If a lymphocyte or other tissue cell could be shown to undergo locomotion without simultaneously accumulating attached materials (even univalent antibodies) into caps, or if cells were shown to cap a multivalent antibody without simultaneously capping a univalent antibody (or lectin) then this hypothesis would be disproved.

How then can one explain the observation that multivalent antibodies do cause capping even on unattached lymphocytes, while univalent ones do not? As a possible explanation, I would suggest that attachment to some sort of surface may be needed for the cell to develop directionality, and to form discrete areas of membrane disassembly and reassembly. In the absence of a solid substratum, the cross linking of the cell surface by multivalent antibodies or lectins may serve as a sort of substitute substratum, stimulating the development of a directional membrane flow and thus capping. This possibility is indirectly supported by the observation of de Petris and Raff[11] that cross linking the lymphocyte surface with multivalent antibodies increases the proportion of cells with discrete uroid tail regions, indicative of cell polarity.

These considerations lead to a pair of experimentally testable predictions regarding the accumulation of univalent antibodies and lectins into caps. The first of these predictions is that even univalent materials should cap if the cells are allowed to attach to a solid substratum and to undergo locomotion. Such caps should form at the trailing margins of the cell. Capping of univalent antibodies on motile polymorphonuclear leukocytes has actually been reported[12] but this experiment does not seem to have been undertaken on lymphocytes. The second prediction is that suspended lymphocytes should cap even univalent antibodies, if their surfaces are simultaneously cross linked by multivalent antibodies or lectins. This should occur even if the multivalent and univalent antibodies are specific for completely different and independent surface molecules. In fact, this prediction can be used to test Bretscher's hypothesis relative to the one that I have proposed above.

According to Bretscher's hypothesis, if multivalent antibodies to one membrane molecule and univalent antibodies to another membrane molecule were simultaneously bound to the surface of the same cell, then the multivalent antibodies should accumulate into a cap, while the univalent antibody should remain dispersed over the cell surface. According to my hypothesis, however, both multivalent and univalent antibodies should accumulate together in the same caps. I would suggest that someone should carry out these experiments.

ALBERT K. HARRIS

Department of Zoology,
University of North Carolina, Chapel Hill,
North Carolina 27514

1 Bretscher, M. S., *Nature*, 260, 21–23 (1976).
2 Taylor, R. B., Duffus, W. P. H., Raff, M. C., and de Petris, S., *Nature new Biol.*, 233, 225–229 (1971).
3 Edidin, M., and Weiss, A., *Proc. natn. Acad. Sci. U.S.A.*, 69, 2456–2459 (1972).
4 Abercrombie, M., Heaysman, J. E. M., and Pegrum, S. M., *Expl Cell Res.*, 62, 389–398 (1970).
5 Harris, A. K., *Acta Protozool.*, 9, 145–151 (1972).
6 Harris, A. K., and Dunn, G. A., *Expl Cell Res.*, 73, 519–523 (1972).
7 Harris, A. K., in *Locomotion of Tissue Cells, Ciba Foundation Symposium 14*, 3–20 (Churchill, London, 1973).
8 Glasstone, S., *Textbook of Physical Chemistry*, 831 (Macmillan, London, 1956).
9 Wirtz, K. W. A., and Zilversmit, D. B., *J. biol. Chem.*, 243, 3596–3602 (1968).
10 McMurray, W. C., and Dawson, R. M. C., *Biochem. J.*, 112, 91–108 (1969).
11 de Petris, S., and Raff, M., in *Locomotion of Tissue Cells, Ciba Foundation Symposium 14*, 27–41 (Churchill, London, 1973).
12 Ryan, G. B., Borysenko, J. Z., and Karnovsky, M. J., *J. Cell Biol.*, 62, 351–365 (1974).

Lack of effect of naloxone on pain perception in humans

RECENT studies have revealed the existence of endogenous morphine-like compounds in the central nervous system[1-3] and in the pituitary[4]. These compounds, named enkephalin[1] and C-fragment[4], have been reported to produce analgesia when injected intraventricularly in rats and cats[5,6] and to be antagonised by naloxone which is a potent antagonist of narcotic analgesics. They also have a high affinity for the opiate receptor, which has been discovered in the central nervous system of vertebrates including man[7], and they can displace naloxone from these receptors[2-4]. There is substantial evidence that central grey structures in the midbrain are intimately involved in analgesia. Electrical stimulation[8] and microinjection of morphine[9,10] in these regions produce analgesia which is antagonised by naloxone. Stimulation of the central grey region in humans[11] has been reported to produce marked analgesia which is antagonised by small doses of naloxone[12]. In addition, acupuncture analgesia has been reported to be antagonised by naloxone[13]. These findings have led to the hypothesis that there may be an ongoing release within the brain of a morphine-like compound which is partially responsible for pain thresholds. If this were so, then the antagonism of the effects of the endogenous compound by naloxone should affect pain perception. The experiments reported here demonstrate that the perception of experimentally induced pain in normal human subjects is not altered by administration of the opiate antagonist naloxone.

Experiments were carried out on five healthy adult subjects. Each subject was tested on three different days for the effects of intravenous injection of 0.4 mg to 0.8 mg naloxone or a saline placebo, which were administered in a double-blind counterbalanced fashion. Pain was induced by delivering incrementally increasing electric shocks starting at and increasing by 0.2 mA at 1.2-s intervals using a Tursky–Watson type[14] concentric electrode attached to the midpoint of the volar surface of the forearm. The following four parameters were obtained: threshold for sensation, pain threshold, severe pain and maximal tolerance current levels. Subjects were in a semi-recumbent position in a sound-proof room. Respiration rate, skin conductance, heart rate, pulse volume and pupil size were also measured and will be reported in a subsequent paper. The experimental schedule was as follows: 10 min resting and basal recording, one pain test trial session, three consecutive pain trials, 10 min relaxation and basal recording, injection of the drug or placebo and repetition of whole sequence without the trial session.

Figure 1 shows an example of the results obtained from one subject on three different days. The small increase in levels observed after 0.8 mg naloxone is not a consistent finding. These results and those obtained from the other subjects reveal that naloxone does not have a significant effect on any of the four parameters measured in this experiment. The average for the five subjects of the ratios of post-drug level (D) were : 1.06 ± 0.06 (s.d.) for placebo, 1.02 ± 0.18 for 0.4 mg naloxone and 1.03 ± 0.1 for 0.8 mg naloxone.

The lack of effect on perception of pain in our studies suggests that in our experimental conditions there is no significant ongoing release of a morphine-like compound. The amount of naloxone used in our experiments should have been more than adequate to have antagonised any endogenous morphine-like compound present. The dose of 0.4 mg is that recommended to antagonise respiratory depression produced by morphine poisoning and acts within a few minutes[15]. Furthermore, 0.1–0.2 mg naloxone has been reported[12] to significantly antagonise analgesia
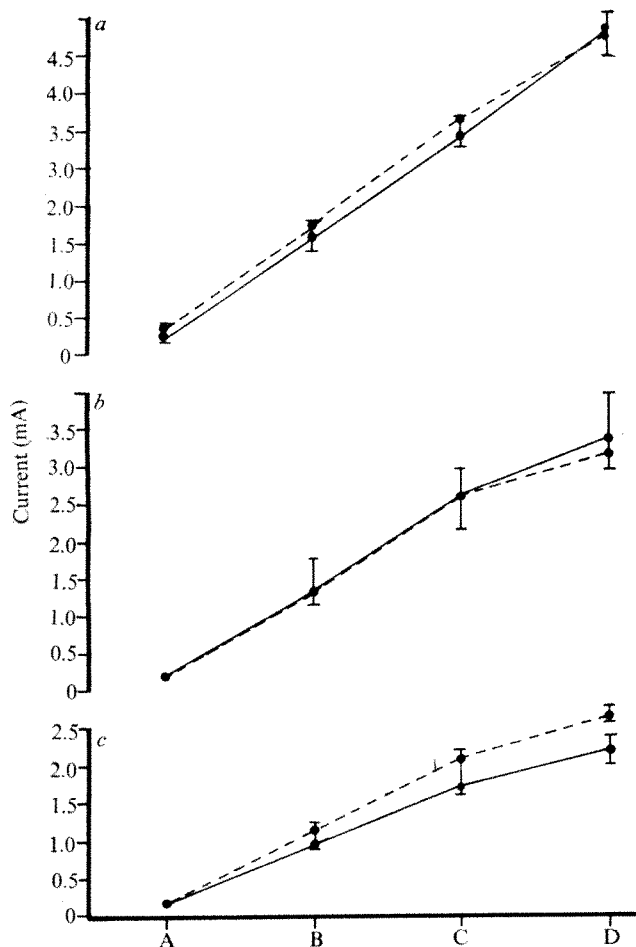
Fig. 1 Results from subject showing lack of effect of placebo (*a*), 0.4 mg naloxone (*b*) or 0.8 mg naloxone (*c*), tested on three different days, on sensation threshold (A), pain threshold (B), severe pain (C), and maximal tolerance (D). Each point is the average of three consecutive trials and the actual range of values for each of the trials is indicated by the vertical 'error' bar; these demonstrate the good repeatability of thresholds from one trial to the next on any one day. Where the ranges of the pre-drug (—) and post-drug (- - -) results overlap, one combined line has been drawn.

is that they are not involved with pain mechanisms. This would be very surprising in view of the mass of experiments on animals showing numerous connections between opiate receptor concentrations, enkephalin distribution and analgesia sensitivity to morphine administration. Another possibility is that the endogenous morphine-like compound is released only under special circumstances. The existence of special psychological states when pain is not perceived even when the person is severely wounded is well documented in war and some accidents[18].

ADEL EL-SOBKY

Neuropsychology Research Unit,
University College Hospital Medical School,
Friern Hospital,
London N11 3BP, UK

JONATHAN O. DOSTROVSKY
PATRICK D. WALL

Cerebral Functions Group,
Department of Anatomy,
University College,
London WC1E 6BT, UK

1 Hughes, J., *et al.*, *Nature*, **258**, 577–579 (1975).
2 Terenius, L., and Wahlström, A., *Acta physiol. scand.*, **94**, 74–81 (1975).
3 Pasternak, G. W., Goodman, R., and Snyder, S. H., *Life Sci.*, **16**, 1765–1769 (1975).
4 Bradbury, A. F., Smyth, D. G., Snell, C. R., Birdsall, N. J. M., and Hulme, E. C., *Nature*, **260**, 793–795 (1976).
5 Beluzzi, J. D., *et al.*, *Nature*, **260**, 625–626 (1976).
6 Feldberg, W., and Smyth, D. G., *J. Physiol., Lond.*, **260**, 30P–31P (1976).
7 Pert, C. B., and Snyder, S. A., *Science*, **179**, 184–188 (1973).
8 Mayer, D. J., and Liebeskind, J. C., *Brain Res.*, **68**, 73–93 (1974).
9 Jacquet, Y. F., and Lajtha, A., *Science*, **182**, 490–492 (1973).
10 Pert, A., and Yaksh, T., *Brain Res.*, **80**, 135–140 (1974).
11 Richardson, D. E., *IEEE Trans. Bio. Med. Eng.*, **23**, 304–306 (1976).
12 Adams, J. E., *Pain*, **2**, 161–166 (1976).
13 Mayer, D. J., Price, D. D., and Rafii, A., *First World Congr. on Pain, Advances in Pain Research and Therapy*, 1 (Raven, New York, in the press).
14 Tursky, B., Watson, P. D., and O'Connell, D. W., *Psychophysiology* **1**, 296–298 (1965).
15 *Winthrop Lab. Data sheet for Narcan* (1974).
16 Parry, L. W., Smith, G. M., and Denton, J. E., *Anesthesia and Analgesia*, **51**, 573–577 (1972).
17 Goldstein, A., Pryor, G. T., Otis, L. S., and Larsen, F., *Life Sci.*, **18**, 599–604 (1976).
18 Melzack, R., *The Puzzle of Pain* (Penguin, Harmondsworth, 1973).

produced by brainstem stimulation in man within 10 min of administration. The method we have used for measuring pain thresholds has been reported reliably to demonstrate the analgesic effects of 10 mg morphine[16]. It is true that morphine is used clinically for long lasting pain rather than our brief experimental pain and it is possible that such pains would be more susceptible to modulation by naloxone and an endogenous narcotic. Also in experiments carried out using the same experimental set up (A. El-Sobky, unpublished) used for this study it was found that administration of heroin increased severe pain and maximum tolerance thresholds in heroin addicts. Therefore, it is unlikely that we would have missed any significant effort of naloxone on pain perception. The lack of effect of naloxone on man's pain thresholds which we report here may agree with Goldstein's[17] failure to find a change in rats' avoidance threshold with naloxone.

Since it is unlikely that there is a tonic release of an endogenous morphine-like compound which interacts with pain perception in normal conditions we are left with the problem of explaining the function of opiate receptors and enkephalin in the central nervous system. One possibility

Determination of acetylcholine null potential in mouse pancreatic acinar cells

PERHAPS the most important parameter of transmitter action is the value of the membrane potential at which no potential change occurs after excitation. This value is termed the transmitter null (equilibrium) potential (refs 1 and 2 and refs therein). The determination of this value requires the use of two intracellular electrodes, one for recording the cell membrane potential and the other for passing current to set the membrane potential at appropriate levels. In addition it requires the availability of reproducible short lasting applications of transmitter. The recent demonstration that the pancreatic acinus functions as one electrical unit, that is that unimpaired electrical communication between neighbouring acinar cells exists[3,4], makes it possible to treat the acinus as one cell. By combining insertion of two electrodes into one acinus with local iontophoretic application of acetylcholine (ACh) from a third extracellular microelectrode, we have measured directly the null potential for the action of a neurotransmitter on mammalian gland cells.

Two separate KCl-filled microelectrodes were inserted very close to each other (20–50 μm intertip distance, viewed under Leitz stereomicroscope (×160 magnification)) into surface acinar cells of mouse pancreas *in vitro* as previously described[3].

The tip of a third microelectrode filled with 2 M AChCl was placed close (< 100 μm) to the impaled acinus. The details of the circuits used for membrane potential measurement, intracellular current injection and iontophoretic ACh application have been described in detail[3,5,6]. The composition of the physiological saline solution flowing through the tissue bath was (mM): NaCl, 103; KCl, 4.7; CaCl$_2$, 2.56; MgCl$_2$, 1.13; NaHCO$_3$, 25; NaH$_2$PO$_4$, 1.15; D-glucose, 2.8; Na pyruvate, 4.9; Na glutamate, 4.9; Na fumarate, 2.7. It was gassed with 95% O$_2$, 5% CO$_2$.

Figure 1 shows the simultaneous recording of membrane potential and resistance from two electrically coupled acinar cells. Iontophoretic ACh applications resulted in reproducible potential and resistance changes of exactly the same size and shape in both cells. If currents (d.c.) of varying size were injected through one of the intracellular microelectrodes, the membrane potential, measured with the other microelectrode, could be set at appropriate levels. Figure 2 shows the marked effect of changes in resting potential on the ACh-evoked potential change. The size of the ACh-evoked depolarisation was clearly reduced as the resting potential was diminished and finally reversed into hyperpolarisation at the very lowest level of membrane potential. In the case represented by Fig. 2, the ACh null potential was – 10 mV. In four other tissues, values ranging from – 5 to – 15 mV were obtained.

In the simplest possible model of transmitter action the membrane potential change (ΔV) after excitation is given by

$$\Delta V(t) = \frac{G_s(t)}{G_s(t) + G_0} (e_x - V_m(t)) \quad \text{where}$$

$G_s(t)$ is conductance of activated synaptic region, G_0 conductance of non-synaptic region, e_x transmitter null potential and $V_m(t)$ membrane potential[1]. If one inserts figures for ΔV, G_s/G_0 and V_m taken from experiments like that shown in Fig. 1 (for example, for t corresponding to time of maximal ACh-evoked depolarisation) values for e_x of about – 3 to – 1 mV are obtained. The slight discrepancy between these values and those determined directly (Fig. 2) is a consequence of the slightly nonlinear relationship between ΔV and resting membrane potential at low resting potentials. The directly determined ACh null potential of – 5 to – 15 mV is rather similar to that in the motor endplate[1] where ACh mainly evokes an increase in Na$^+$ and K$^+$ conductance. The ACh-evoked potential and resistance change in pancreatic acinar cells is very sensitive to reductions in extracellular Na concentration[5] and it has been concluded tentatively that the mechanism of action of ACh in the pancreatic acinar cells is to increase Na$^+$ and K$^+$ conductance[1]. The experiments reported here conclusively demonstrate

Fig. 1 Effect of short pulses of iontophoretic ACh application on membrane potential and resistance in two electrically coupled acinar cells. The top part shows the pen recordings, and below is an oscilloscope photograph consisting of two exposures at times marked a and b in the pen recordings. The interruption of the pen recordings represents a 5-min interval. The microelectrode measuring the potential represented in the lower part of the pen recording was inserted first. Subsequently a second microelectrode, through which square wave current pulses could be injected, was inserted. The potential measured by this electrode is shown in the upper part of the pen recording. The short lasting current injections gave rise to short lasting membrane hyperpolarisations, both in the cell into which the current was injected (upper trace) (a circuit was used to compensate for potential changes caused by the current passage through the electrode tip resistance[6], in the beginning of the record the tip resistance compensation is imperfect and is corrected) and in the neighbouring cell (lower trace). The time course of the polarisations is seen in the oscilloscope photograph (calibration: horizontal 20 ms, vertical 10 mV and 2 × 10^{-9} A). The coupling ratio (current pulse-induced potential change in the cell of current injection divided by electrotonic potential change in neighbouring cell) is 1. In the event marker trace in the middle are shown signals marking the periods of ejecting current through the extracellular AChCl electrode. A retaining d.c. current of 2 × 10^{-8} A, to prevent spontaneous leakage of ACh, was used and the ejecting current pulses were of 0.5 s duration and 6 × 10^{-8} A strength.
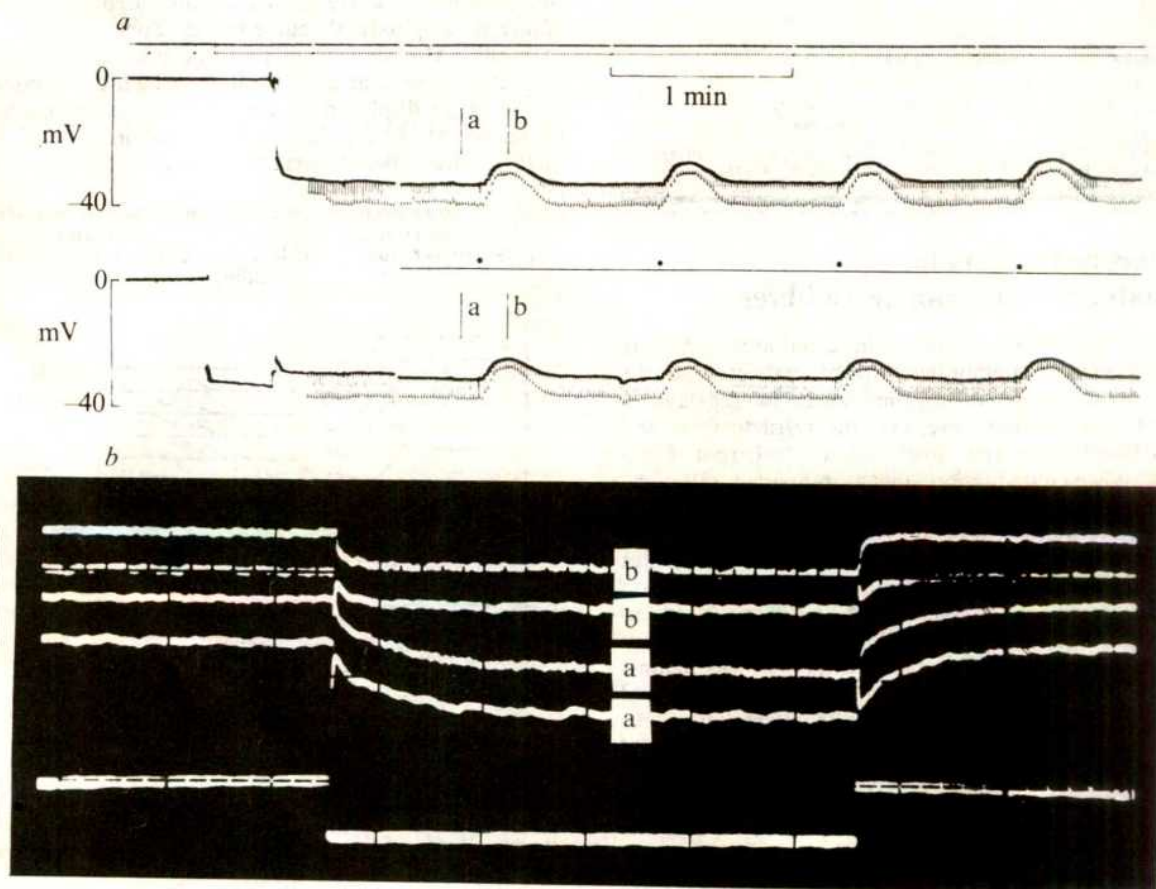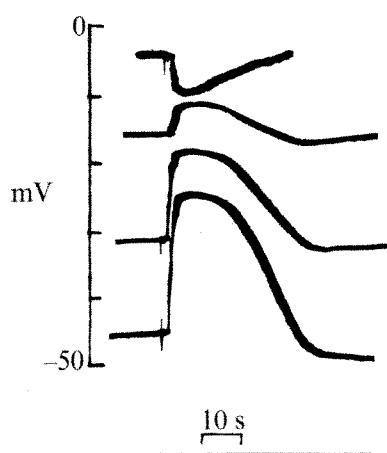
Fig. 2 Effect of short pulses (constant size and duration) of iontophoretic ACh application on membrane potential recorded at four different values of resting membrane potential. The spontaneous resting potential was −31 mV. Higher and lower resting potentials were obtained by passing d.c. hyperpolarising or depolarising currents, respectively, through a microelectrode inserted into a neighbouring cell. A d.c. retaining current of 2×10^{-8}A was applied to the extracellular AChCl electrode and ejecting current pulses of 0.5 s duration and 6×10^{-8}A size were applied at time of marker signal in bottom trace.

that ACh acts on the pancreatic acinar cell membrane by increasing ion conductance and are consistent with the hypothesis that the ion species involved are Na^+ and K^+.

We thank Mrs M. Gordon for technical assistance and the MRC for grant support.

N. IWATSUKI
O. H. PETERSEN

Department of Physiology,
University of Dundee,
Dundee DD1 4HN, UK

1 Gage, P. W., *Physiol. Rev.*, **56**, 177–247 (1976).
2 Ginsborg, B. L., *Biochim. biophys. Acta*, **300**, 289–317 (1973).
3 Petersen, O. H., and Ueda, N., *J. Physiol., Lond.*, **254**, 583–606 (1976).
4 Petersen, O. H., *Physiol. Rev.*, **56**, 535–577 (1976).
5 Nishiyama, A., and Petersen, O. H., *J. Physiol., Lond.*, **244**, 431–465 (1975).
6 Nishiyama, A., and Petersen, O. H., *J. Physiol., Lond.*, **238**, 145–158 (1974).

Continuous conduction in demyelinated mammalian nerve fibres

MYELINATED nerve fibres conduct in a saltatory fashion, with sites of inward membrane current restricted to the nodes of Ranvier[1]. Loss of myelin causes long delays in the internodal conduction time, extreme refractoriness and conduction block[2]. In the large rat ventral root fibres studied the slowed conduction always remained saltatory, with discrete sites of inward current, but the limited spatial resolution of the method used prevented conclusions about whether any membrane beyond that of the original nodes was excited. To find out more about the internodal axon membrane exposed by demyelination, we have used an improved technique of external longitudinal current analysis. A smaller electrode separation (120 μm as against 400–600 μm) has improved the spatial resolution[3], and signal averaging has made possible recordings from smaller axons. We have found that in these fibres not only is internodal axon membrane electrically excitable, but that it can support continuous conduction across a demyelinated internode. This result provides a new basis for interpreting the effects of demyelinating disease.

Longitudinal current records from a normal fibre together with their relationships to three nodes are shown in Fig. 1;
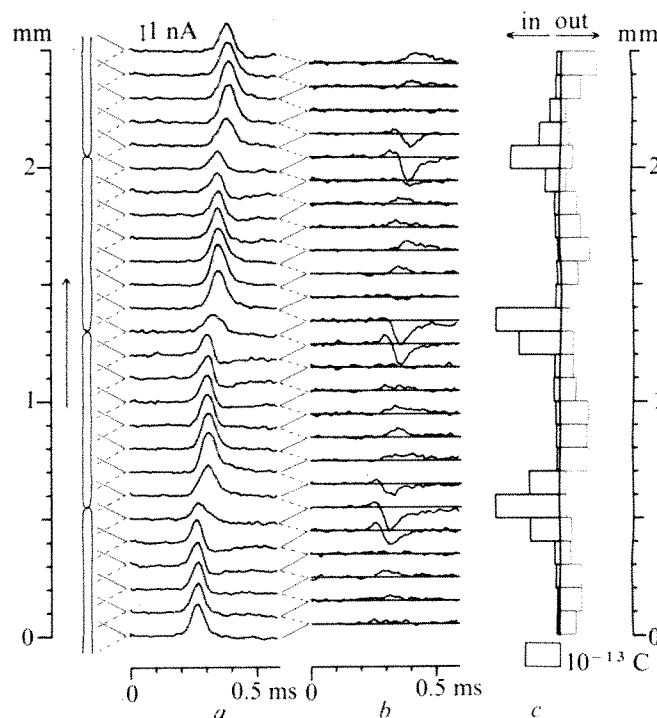
Fig. 1 Action currents in normal ventral root fibre, showing saltatory conduction. *a*, Average longitudinal currents (1,024 sweeps) recorded from two platinum wires, 120 μm apart, which were slid 100 μm along the root between averages. The approximate positions of the electrodes with respect to the single active fibre are indicated on the left, with the sites of the nodes of Ranvier inferred from the records. The arrow indicates the direction of propagation of the impulse. *b*, Membrane currents (inward current downwards) obtained by subtracting longitudinal currents as indicated. *c*, Inward (thick line) and outward (thin line) current integrals calculated from (*b*).

the membrane currents were calculated by differencing adjacent longitudinal currents as done by Huxley and Stampfli[1]. Also shown are the time integrals of inward and outward currents at each position, which correspond to the total charge displaced in either direction for each 100 μm of nerve. These records confirm that in normal rat myelinated fibres inward current is confined to the nodes, and

Fig. 2 Action currents in ventral root fibre partially demyelinated with diphtheria toxin, showing stretch of continuous conduction. Explanations of (*a*), (*b*) and (*c*) are as for Fig. 1, but with different scaling.
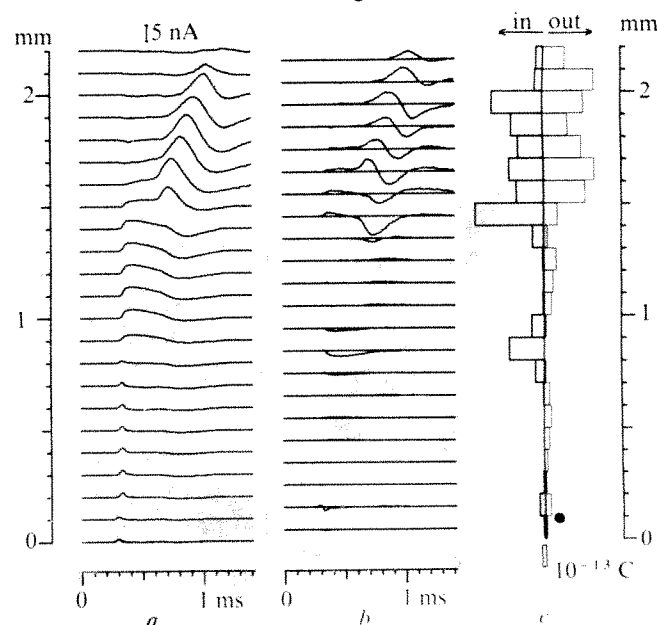
provides a control for comparison with demyelinated fibres. Figures 2 and 3 are corresponding records for two fibres from animals receiving intrathecal injections of diphtheria toxin 6 d previously, in which the action potentials were propagated in a continuous fashion for about 500 μm. This conduction is termed continuous, because (1), there is a continuous change in latency of the longitudinal currents from one position to the next, in contrast to normal saltatory latency shift, and (2) each 100 μm of nerve in turn becomes a source of inward current, indicating a continuous progression of excitation. The time integrals of inward and outward current reveal that the propagation though continuous, is not even. They also show that at least as much charge crosses the axon membrane in the inward direction along the newly bared internode as at the adjacent nodes.
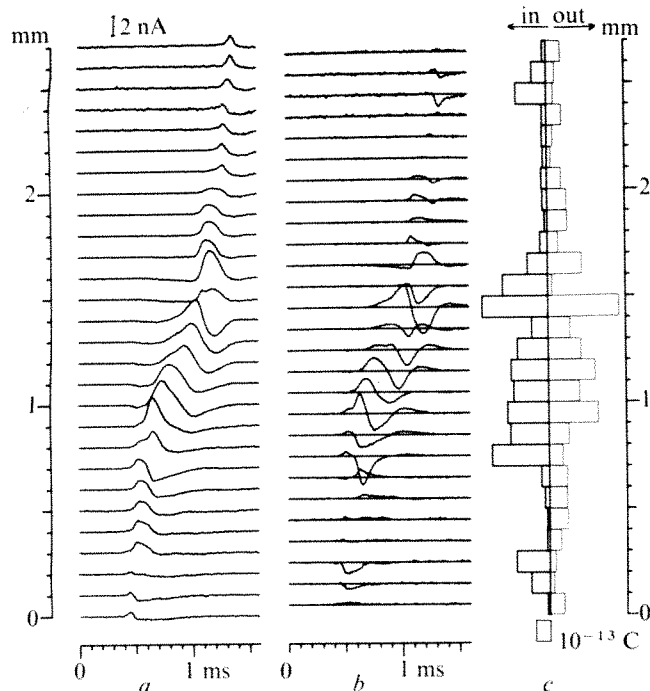
Fig. 3 Action currents in partially demyelinated fibre, showing stretch of continuous conduction and resumption of saltatory conduction. Explanations of (*a*), (*b*) and (*c*) are as for Fig. 1, but with different scaling.

We have recorded five examples of continuous conduction in fibres with internodal spacings in the range 500–850 μm. In three cases (including the two illustrated) the length of continuous conduction was about 500 μm, and the velocity corresponded to roughly 1/20th of the velocity expected for normal stretches of the same fibre. This figure is similar to the ratio of velocity of continuous conduction in the amyelinated roots of dystrophic mice to normal controls[4,5]. The remaining two cases were of conduction continuous over distances less than 300 μm. Our failure so far to observe such continuous conduction in demyelinated fibres with internodal distances greater than 850 μm (corresponding to an external fibre diameter of about 6 μm), and the previous failure to observe the phenomenon at all[2], indicates some difference between larger and smaller nerves. It is by no means clear, however, that internodal excitability varies with fibre size, since differences in pathology or geometry could also account for the restriction of continuous conduction to small fibres in our experiments.

H. BOSTOCK
T. A. SEARS

*Sobell Department of Neurophysiology,
Institute of Neurology,
Queen Square,
London WC1 3BG, UK*

1 Huxley, A. F., and Stampfli, R., *J. Physiol., Lond.*, 108, 315–339 (1949).
2 Rasminsky, M., and Sears, T. A., *J. Physiol., Lond.*, 227, 323–350 (1972).
3 Bostock, H., and Sears, T. A., *J. Physiol., Lond.*, 252, 72–73P (1975).
4 Huizar, P., Kuno, M., and Miyata, Y., *J. Physiol., Lond.*, 248, 231–246 (1975).
5 Rasminsky, M., and Kearney, R. E., *Neurology*, 26, 367 (1976).

Epstein–Barr virus genome in infectious mononucleosis

THE seroepidemiological evidence[1] linking the Epstein–Barr virus (EBV) causatively with infectious mononucleosis is compelling, but there are many questions concerning the precise nature of the relationship between virus and disease[2] and the immunological responses that limit the disease[3,4]. For example, until recently it remained unproven that the transforming agent excreted from the throat during acute and convalescent infectious mononucleosis is EBV specific. Lymphocytes from the umbilical cord can be transformed into continuous cell lines by exposure to the virus excreted from the throat. New evidence presented here and data published recently show that such transformed lymphocytes contain EBV genetic material as well as a virus-specific nuclear protein[5] even though such cells, unlike those isolated from the peripheral blood, do not shed virus. Another question, persistent and crucial, has been whether the EBV originally derived from Burkitt's lymphoma in Uganda and the EBV associated with infectious mononucleosis are identical. We show here that the viral DNA in human lymphocytic lines derived from infectious mononucleosis partially lacks homology to EBV DNA. Such differences imply that an extensively deleted, defective EBV genome is retained in lymphocyte lines established from infectious mononucleosis or that the cell lines harbour different strains of EBV with up to 35% heterologous DNA sequences.

Students in the infirmary of the University of North Carolina at Chapel Hill were diagnosed as having mononucleosis on the basis of fever, lymphadenopathy, sore throat, splenic enlargement, atypical lymphocytosis and a positive Monospot test (Ortho Diagnostics). The diagnosis was confirmed by the presence of antibodies to EBV capsid antigen (VCA) and to early antigen (EA) as well as by the other data presented in this communication.

Complementary RNA–DNA hybridisation was carried out on duplicate membranes on which had been immobilised 20–80 μg of the cellular DNA to be analysed. Tritiated EBV-specific complementary RNA (cRNA) that was synthesised *in vitro* on a template of EBV DNA with DNA-dependent RNA polymerase from *Escherichia coli* was hybridised to the denatured cellular DNA as described before[6]. The specificity of the ^{3}H-EBV cRNA and the sensitivity of the test have been defined; as few as two EBV genome equivalents per cell can be detected.

DNA–DNA renaturation kinetics analyses were carried out with EBV DNA tritiated *in vitro* as previously described[7]. The reassociated DNA was distinguished from unrenatured DNA by digestion with S1 single-strand-specific exonuclease derived from *Aspergillus oryzae* as described before[6]. We express results as C_0t values (mol s l^{-1}) in which C_0 is the concentration of double-stranded DNA; estimates of EBV genome numbers were based on the $C_0t_{\frac{1}{2}}$ values[6].

In Table 1 are some of the data from a series of patients with proven mononucleosis. Lymphoblastoid cell lines were established from the peripheral blood of the patients after 3–4 weeks of culture *in vitro*. In addition to EBV capsid and early antigens, each of these cell lines contained EBV DNA as shown by cRNA–DNA hybridisation.

Filtered throat washings from the same patients when applied to umbilical cord leukocytes regularly led to the establishment of lymphoblastoid cell lines. Cord lymphocytes fail to proliferate if they are not exposed to the throat washings[8]. These transformed cells do not contain capsid or early antigens (refs 10 and 11 and Table 1). The hybridisation data show that there is EBV genetic material in the cell lines which suggests the

Table 1 Lymphoid cell lines derived from the peripheral blood of patients with infectious mononucleosis or from cord blood leukocytes transformed with throat washings

Patient	EBV anti-body titres VCA	EA	Cell lines from peripheral blood No. of genome equivalents per cell*	% Cells with VCA	% Cells with EA	Cord blood leukocytes transformed with throat washings No. of genome equivalents per cell*	% Cells with VCA	% Cells with EA
LH	40	ND	53	1.1	1.7	ND	ND	ND
DT	160	160	20	0.4	1.2	15	0†	0
MB	320	80	37	0.7	1.1	12	0	0
NW	160	160	29	0.7	1.0	11	0	0
KB	320	160	39	0.9	1.3	11	0	0
EB	160	160	34	0.8	1.5	10	0	0

*Determined by RNA–DNA hybridisation with ^{3}H-EBV-specific cRNA as probe. The specific activity of the cRNA was 6×10^6 c.p.m. µg^{-1}; an input of 10^5 c.p.m. of cRNA was applied to duplicate samples of 20–50 µg of cellular DNA as described before[6].
†Less than 1 in 10^5.
ND, not done. Reciprocals of serum dilutions are given.

virus-specific nature of the transformation. Such cells also contain EBV nuclear antigen (EBNA)[10]. The relative number of genome equivalents found in the lines is consistently less than that found in peripheral blood cell lines. Explanted peripheral blood cells sometimes contain large numbers of EBV genomes

Fig. 1 Incomplete homology to the EBV genome of the viral DNA sequences in lymphocyte lines from infectious mononucleosis. Determinations were made by DNA–DNA renaturation kinetics analysis. *a*, ×, ^{3}H-EBV DNA (0.02 µg) + calf thymus DNA (2 mg); ○, PB20 DNA (1.1 mg) was from a cell line derived from peripheral blood; ●, TW20 DNA (1.8 mg) was a cord lymphocyte line transformed by throat washings from the same patient; □, HR1K DNA (2 mg) was homologous DNA from the EBV-producing cell line that was the source of the index EBV DNA. *b*: △, PB16 DNA (1.5 mg) was from a cell line established from the peripheral blood of another patient with infectious mononucleosis; ▲, TW16 DNA (1.5 mg) was from a cell line produced by exposure to throat washings from the same patient. Other symbols as in (*a*). Labelled EBV DNA (0.02 µg, specific activity 1.8×10^6 c.p.m. µg^{-1}) was present in all renaturation mixtures. The total concentration of DNA in all mixtures was brought to 2 mg by addition of calf thymus DNA before denaturation and reannealing.

and may even replicate virus[12]; this does not occur with the transformed cord cell lines.

Finally, virus shed from one of the peripheral blood cell lines (PB6) could transform cord blood lymphocytes. The secondarily transformed cells also contained a relatively small amount of EBV DNA, but no detectable capsid or early antigens. This last finding indicates that peripheral lymphocytes when transformed into cell lines can shed biologically active EBV.

Perhaps not all the DNA sequences found in the genome of the standard lymphoma-derived P3HR1 virus[13] exist in the mononucleosis cell lines. Another possibility is that the cord blood lymphocytes, which contain fewer copies of viral DNA and do not express viral structural antigens, do not contain as full a representation of the EBV genome as the cell lines derived from peripheral blood.

Figure 1 shows an analysis of the DNA extracted from pairs of lymphoid cell lines, one from peripheral blood (PB), and one from cord blood cells that had been transformed by throat washings (TW). Each pair of cell lines, PB and TW20 and PB and TW16, derive from materials from a single patient. We found that neither of the peripheral blood cell lines contained all the viral DNA sequences of the prototype EBV genome. The PB20 cells lacked 30–35% of the homologous viral DNA, and the PB16 line lacked 25–30% of the homologous sequences found in the index viral DNA. Nevertheless, both of these cell lines did produce EBV capsid antigens and early antigens. This means either that the whole of the EBV genome is not needed for the coding and translation of these antigens, and that the genome is defective, or that full or unit-length viral DNA is present in the two lines, but that 25–35% of this viral genome contains heterologous DNA sequences.

Another conclusion from these data is that cord blood lymphocytes transformed by the throat washings contain approximately the same viral DNA sequences as are found in the cell lines derived from peripheral blood. Apparently the cord blood lymphocytes do not necessarily select for defective genome. The reason for the sharply restricted expression of the genome contained in the cord cell lines is therefore not apparent.

Cord cells transformed by cell-free extracts from a peripheral blood cell line (PB6) contained about 80% of the DNA sequences found in the index viral genome (Fig. 2*a*).

Figure 2*a* shows the results of analysis of the DNA of EBV-transformed marmoset cells. This line originated from marmoset lymphocytes that had been transformed by throat washings obtained from a patient with infectious mononucleosis[14]. Here approximately 40% of the DNA sequences found in the P3HR1-derived virus are not present in the marmoset line.

An analysis of Raji DNA is also shown (Fig. 2*b*). The completeness of reannealing of the homologous DNA approaches 100%. This is expected as Raji cells are thought to contain unit-length viral DNA[15] and can be induced with IUdR to make virus[16].

Understanding of the pathogenesis of infectious mono-

Fig. 2 Extent of homology of the viral DNA sequences found in various lymphoblastoid cell lines. *a*, △, B95–8 DNA (2 mg) was from a marmoset cell line transformed by human EBV; ▽, PB6 DNA (1.3 mg) was from a human cord cell line transformed by virus from a cell line established from the peripheral blood of a patient with infectious mononucleosis; ×, ^{3}H-EBV DNA; ■, Raji DNA; □, HR1K DNA. *b*, Three concentrations of Raji DNA (1, 0.42 mg; 2, 0.56 mg; 3, 0.84 mg), a human lymphoblastoid cell line derived from an African Burkitt's lymphoma; these cells contained fully homologous EBV DNA. ●, ^{3}H-EBV DNA; ○, HR1K DNA (2 mg).

nucleosis is still deficient. The virus apparently replicates somewhere in the oropharynx, including perhaps the parotid gland[17]. The precise site and cell type in which the virus replicates, however, are not known.

Earlier analyses[8] suggested that all the P3HR1 EBV DNA sequences were present in three different IM cell lines, but our results obtained with different cell lines show clear differences. Up to 35% of the EBV DNA sequences were missing in cell lines established from peripheral blood as well as in cord blood cells transformed with fresh virus isolates. We do not know whether the molecular weight of the homologous DNA in the cell lines was 35% less than that of virion DNA. Defective genomes might be more liable to persist in lymphocytes, although clearly some lymphocytes contain replicative amounts of genome; explantation of peripheral blood cells can lead to shedding of virus.

In the case of the B95-8 virus, a fuller analysis has been possible. Pritchett *et al.*[18] showed that this virus lacked at least 15% of the DNA sequences found in the P3HR1 EBV, yet the molecular weight of the B95-8 virus DNA was the same as that of EBV. The missing sequences were thought to be made up of repetitions of DNA sequences found in EBV; no heterologous sequences were detected.

Are there strain variants of EBV? The missing sequences in EBV associated with infectious mononucleosis may be replaced by heterologous DNA sequences so that there are different strains of EBV bearing extensive but not complete homology to the EBV derived from an African lymphoma. Epidemiologi-

cally and pathogenetically this is an attractive possibility as brought out recently by Miller *et al.*[11], but we cannot distinguish such heterologous viral DNA sequences because of the lack of any other EBV strain in the quantity needed for reciprocal homology studies.

Nasopharyngeal carcinomas also contain viral DNA only partly homologous to EBV[19]. Recent analysis of herpes simplex type 1 and type 2 DNA by restriction endonuclease digestion and similar analyses of cytomegalovirus DNA demonstrate the existence of numerous strains of these viruses hitherto undefined[20-22]. The biological significance of this diversity is undetermined. In the case of herpes simplex virus both defective genomes and strain variants exist, and we believe this will prove to be the case with EBV, but we are not yet ready to refer to the Epstein–Barr "viruses".

We thank Drs James McCutchan and James Taylor of the Student Infirmary of the University of North Carolina for making this work possible. We also thank the students who participated and the infirmary physicians who cooperated. We thank Drs Lindsey Hutt, Harry Dascomb and Stanley Lemon for helping in the study of some patients. This research was supported by the John A. Hartford Foundation, Inc.

JOSEPH S. PAGANO
CHIEN-HUI HUANG

Cancer Research Centre,
University of North Carolina,
Chapel Hill, North Carolina 27514

YUNG-TSEN HUANG

Virology Laboratory,
North Carolina Memorial Hospital,
Chapel Hill, North Carolina 27514

1 Niederman, J. C., McCollum, R. W., Henle, G., and Henle, W., *J. Am. med. Assoc.*, 203, 205–209 (1968).
2 Pagano, J. S., in *Viruses, Evolution and Cancer* (edit. by Kurstak, E. and Maramorosch, K.), 79–116 (Academic, New York, 1974).
3 Pagano, J. S., *Cold Spring Harb. Symp. quant. Biol.*, 39, 797–905 (1975).
4 Hutt, L. M., Huang, Y. T., Dascomb, H. E., and Pagano, J. S., *J. Immun.*, 115, 243–248 (1975).
5 Reedman, B. M., and Klein, G., *Int. J. Cancer*, 11, 499–520 (1973).
6 Nonoyama, M., and Pagano, J. S., *Nature new Biol.*, 233, 103–106 (1971).
7 Nonoyama, M., and Pagano, J. S., *Nature*, 242, 44–47 (1973).
8 Kawai, Y., Nonoyama, M., and Pagano, J. S., *J. Virol.*, 12, 1006–1012 (1973).
9 Nilsson, K., Klein, G., Henle, W., and Henle, G., *Int. J. Cancer*, 8, 443–450 (1971).
10 Gerber, P., Nkrumah, F. K., Pritchett, R., and Kieff, E., *Int. J. Cancer*, 17, 71–81 (1976).
11 Miller, G., Coope, D., Niederman, J., and Pagano, J. S., *J. Virol.*, 18, 1071–1080 (1976).
12 Minowada, J., Nonoyama, M., Moore, G. E., Rauch, A. M., and Pagano, J. S., *Cancer Res.*, 34, 1898–1903 (1974).
13 Hinuma, U., *et al.*, *J. Virol.*, 1, 1045 (1967).
14 Miller, G., Shope, T., Lisco, H., Stitt, D., and Lipman, M., *Proc. natn. Acad. Sci. U.S.A.*, 69, 383–387 (1972).
15 Anderson, M., and Lindahl, T., *Virology* (in the press).
16 Hampar, B., Derge, J. G., Martos, L. M., and Walker, J. L., *Proc. natn. Acad. Sci. U.S.A.*, 69, 78–82 (1972).
17 Niederman, J. S., Miller, G., Pearson, H. A., Pagano, J. S., and Dowaliby, J. M., *New Eng. J. Med.*, 294, 1355–1359 (1976).
18 Pritchett, R. F., Hayward, S. D., and Kieff, E. D., *J. Virol.*, 15, 556–569 (1975).
19 Pagano, J. S., *et al.*, in *Proc. second int. Symp. Oncogenesis and Herpesviruses II* (edit. by de Thé, G., Epstein, M. A., and zur Hausen, H.), 179–190 (IARC, Lyon, France, 1975).
20 Skare, J., Summers, W. P., and Summers, W. C., *J. Virol.*, 15, 726 (1975).
21 Hayward, G. E., Frenkel, N., and Roizman, B., *Proc. natn. Acad. Sci. U.S.A.*, 72, 1768–1772 (1975).
22 Kilpatrick, B., Huang, E.-S., and Pagano, J. S., *J. Virol.*, 18, 1095–1105 (1976).

Photoreceptor membrane carbohydrate on the intradiscal surface of retinal rod disks

RHODOPSIN is an integral constituent of the photoreceptor membrane where it is believed to function as a light-activated gate for the release of ions or other substances. The shape, localisation and orientation of the rhodopsin molecule in the membrane is therefore of great importance in understanding the primary events of light perception. Uncertainties in the present state of knowledge are reflected in the numerous models which localise rhodopsin variously on the internal, external or both sides of the disk membrane. In the last few years it has become increasingly accepted that rhodopsin is an elongated molecule, partially embedded in or entirely spanning the disk membrane. Even this model leaves many uncertainties
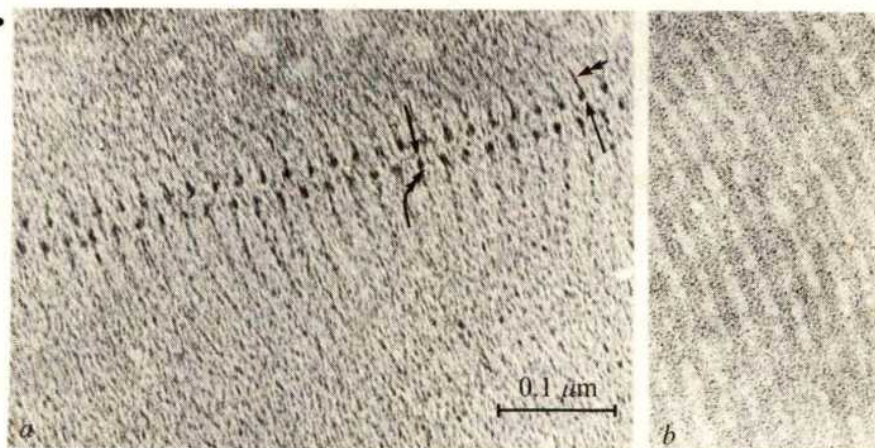
Fig. 1 *a*, Detail of a rod outer segment from a formalin–glutaraldehyde-fixed and Vestopal-embedded frog retina. Carbohydrate reaction (periodic acid-thiosemicarbazide-silver proteinate[11]) was carried out on thin sections. The dense reaction product is seen inside the membrane loops of the disk margins (dark dots, arrows) whence it can be traced as a thin dense line into the interior of the disks delineating the contact (intradiscal) surface of the membranes (double arrow). *b*, Control section from the same material without oxidation with periodic acid. No reaction can be seen inside the disks.

as to the orientation of the molecule in the membrane. Besides the chromophore retinal, the oligosaccharide chain[1,2] is a good orientation point on the rhodopsin molecule. Since rhodopsin is the dominant protein of the photoreceptor membrane[3–8] where it forms about 85–95% of the total protein content, localisation of the carbohydrate component of the membrane would indicate the location of the rhodopsin carbohydrate moiety. Here I present an electron microscopic histochemical demonstration of the carbohydrate component of the photoreceptor membrane; a part of this investigation was reported briefly earlier[9].

In the first group, aldehyde-fixed retinae of frogs (*Rana esculenta*, *Rana pipiens*) were embedded in Epon, Vestopal or glycol methacrylate and electron microscopic histochemical reactions for detection of carbohydrates were carried out on thin sections. Three reactions were based on oxidation of vicinal glycols to form aldehyde groups that were subsequently demonstrated by silver methenamine[10], thiosemicarbazide-

silver proteinate[11] or alkaline bismuth[12]. Oxidation was omitted in the controls. In addition, the carbohydrate staining by phosphotungstic acid at low *p*H (refs 13 and 14) was used because of the better resolving power of the method. All electron microscopic histochemical reactions, from which a representative example is shown on Fig. 1*a*, demonstrated a fine reactive line at the contact (internal) surface of disk membranes. The reaction was especially obvious at the disk margin where it was localised at the internal surface of the membrane loop. The reaction was absent in the control specimens (Fig. 1*b*).

In the second part of this work, ferritin-labelled concanavalin A (con A–Ft) was used to detect lectin-binding carbohydrates on the photoreceptor membrane. The con A–Ft was prepared according to de Petris *et al.*[15] and was a gift of Dr D. Lawson (University College, London). A major difficulty in lectin binding was the inability of the voluminous conjugate to penetrate across plasma and disk membranes that was necessary to expose both sides of the photoreceptor membranes to the lectin. The problem was largely solved by a freeze–thawing method as detailed in the following:

Fig. 2 Ferritin-labelled concanavalin A (con A–Ft) is bound to the internal surface of disk-derived vacuoles. No conjugate can be found at the external surface of the vacuoles. This detail represents a portion of an outer segment from a frog retina that was swollen in ammonium acetate, fixed, freeze-fractured and exposed to con A–Ft.

isolation of frog retinae
↓
swelling in 20 mM ammonium acetate (*p*H 6.9, 4 °C, 15 min)
↓
fixation in 0.5% glutaraldehyde (in 10 mM cacodylate buffer *p*H 7, 4 °C, several hours)
↓
washing in buffer (20 mM cacodylate, 4 °C, 20 h)
↓
freezing in liquid nitrogen and breaking the retinae into little fragments
↓

thawing in con A–Ft solution   thawing in buffer (4 °C, 15 min)
↓   ↓
washing in buffer   20% glycerol in buffer (4 °C, 2 h)
↓   ↓
fixation in osmium   freeze-fracturing to check the sidedness of the membranes
↓
embedding for electron microscopy

Freeze-fracturing was made in a modified freeze-fracturing device (P. R. Környei and Balogh, to be published) that was based on the cold block principle.

The unilateral localisation of carbohydrates was even more unambiguous with the con A–Ft binding than with the histochemical reactions. The swollen disks appeared as more-or-less

inflated vacuoles. Those disk-derived vacuoles which became patent because of the freezing, fracturing and thawing, contained the con A–Ft bound to the internal (intradiscal) surface of the disk membranes (Figs 2 and 4). No con A–Ft particles were ever found associated with the external (interdiscal, cytoplasmic) surface of the disks, although it was freely accessible to the conjugate because of the broken plasma membrane. The latter showed con A–Ft particles bound to its extracellular surface. No binding was observed using α-methyl-D-glucoside together with the lectin.

It is crucial for evaluating the con A-binding experiment whether the original sidedness of the disk membranes has been retained or not. One way to check this is the known asymmetrical distribution of intramembrane particles on freeze-fractured photoreceptor membranes. As can be seen in Fig. 3, the sidedness of the disk membranes was not altered during the preparative procedure since, similarly to intact disk membranes,

Fig. 4 The same material as Fig. 2 to show persisting disk edges in the con A–Ft experiment (arrows). The membrane loops of the disk edges are directly continuous with the interior of the disk-derived vacuoles indicating the original sideness of the disk membranes.
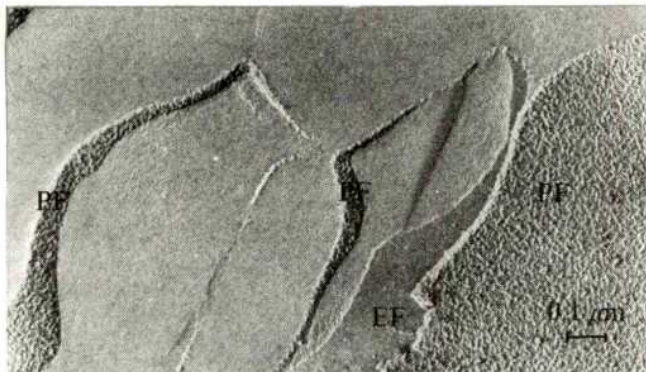
Fig. 3 Freeze-fracture image of swollen disks, treated similarly to those in Fig. 2 (without con A–Ft). The presence of numerous intramembrane particles on fracture face looking towards the lumen of the vacuoles (fracture face of "cytoplasmic" membrane half, PF) indicates that the original sidedness of the disk membranes has not been changed during the preparative procedure. EF, fracture face of the endoplasmic membrane half.

the majority of intramembrane particles was found on fracture face looking towards the interior of the vacuole ("cytoplasmic" membrane half, fracture face PF[16]). Another means of recognising the sidedness of the membranes is the known rigidity of the disk edge which preserves its form even after drastic treatments[17,18]. In our con A–Ft material, disk edges were often observed and their concavity was continuous with the internal space of the disk-derived vacuoles (Fig. 4, arrows). This observation indicates that the disks were only swollen and not everted during the preparative procedure.

Both the carbohydrate reactions on ultrathin sections and the con A–Ft binding on freeze-thawed photoreceptor outer segments showed the presence of a carbohydrate layer at the intradiscal surface of frog retinal rods. The extraordinary molecular simplicity of photoreceptor membranes makes it easy to find the most probable candidate for the carbohydrate layer. Rhodopsin is known to be the dominant protein in outer segment membranes[3-8] (about 9/10 of all membrane proteins being rhodopsin). It is a glycoprotein[1,2] and binds concanavalin A strongly and specifically[19,20]. There can be little doubt therefore that of the proteins present in rod outer segment membranes, rhodopsin is primarily responsible for the lectin binding. The contribution of membrane glycolipids to the carbohydrate reaction seems to be negligible, partly because of the low percentage of glycolipids when compared with the rhodopsin content and partly because an intradiscal carbohydrate reaction was present also in our embedded material. But even if we assume a small contribution of glycolipids to concanavalin-binding, it does not make
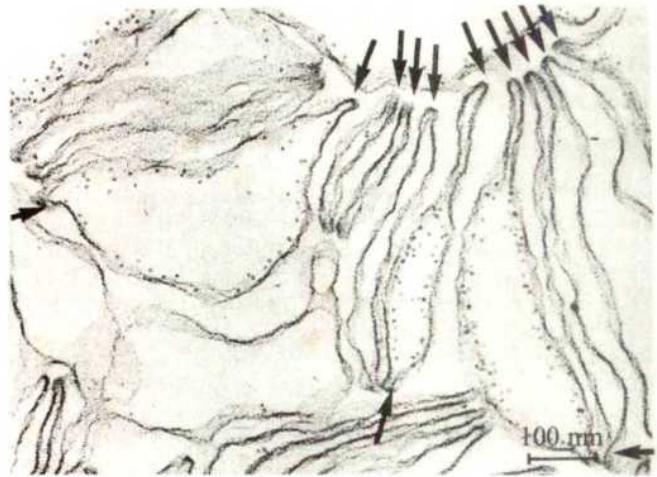
localisation of the rhodopsin carbohydrate moiety on the internal disk surface doubtful due to the unilateral localisation of the lectin-binding sites.

Assuming that rhodopsin is responsible for the carbohydrate reaction, the carbohydrate component of rhodopsin must be localised at the intradiscal surface of the disk membrane. This means that the rhodopsin molecule is either situated in the internal half of the membrane or spans the membrane with its extended oligosaccharide chain[21] protruding into the intradiscal space. Such a localisation of the rhodopsin carbohydrate moiety is supported by observations in which it was shown that the carbohydrate component and the chromophore were not accessible to protease digestion in photoreceptor vesicles, although more than one third of the rhodopsin molecule was digested[22,23]. At present, it is difficult to compare the present findings with investigations[20,24] reporting concanavalin binding on the external surface of isolated disks. To correlate the contradicting results it would be necessary to know whether the photoreceptor vesicles in the cited experiments were completely closed and whether all vesicles showed the original sidedness.

PAL RÖHLICH

Laboratory I of Electron Microscopy, and
 2nd Department of Anatomy, Histology and Embryology,
Semmelweis University of Medicine, 1094 Budapest,
Tuzolto u. 58, Hungary

1 Heller, J., *Biochemistry*, **8**, 675–678 (1969).
2 Heller, J., and Lawrence, M. A., *Biochemistry*, **9**, 864–869 (1970).
3 Bownds, D., Gaide-Huguenin, A. C., *Nature*, **225**, 870–872 (1970).
4 Bownds, D., Gordon-Walker, A., Gaide-Huguenin, A. C., and Robinson, W. J., *J. gen. Physiol.*, **58**, 225–237 (1971).
5 Daemen, J. F. M., DeGrip, W. J., and Jansen, P. A. A., *Biochim. biophys. Acta*, **271**, 419–428 (1972).
6 Hall, M. O., Bok, D., and Bacharach, A. D. E., *J. molec. Biol.*, **45**, 397–406 (1969).
7 Heitzmann, H., *Nature*, **235**, 114 (1972).
8 Papermaster, D. S., and Dreyer, W. J., *Biochemistry*, **13**, 2438–2444 (1974).
9 Röhlich, P., *Third International Meeting of International Society of Neurochemistry Budapest*, Abstr. p. 292 (1971).
10 Rambourg, A., Hernandez, W., and Leblond, C. P., *J. Cell Biol.*, **40**, 395–414 (1969).
11 Thiéry, J. P., *J. Microscopie*, **6**, 987–1018 (1967).
12 Ainsworth, S. K., Ito, S., and Karnovsky, M. J., *J. Histochem. Cytochem.*, **20**, 995–1005 (1972).
13 Rambourg, A., *Electron Microscopy, 1968*, **2** (edit. by Bocciarelli, D. S.), 57–58 (1968).
14 Marinozzi, V., *Electron Microscopy 1968*, **2** (edit. by Bocciarelli, D. S.), 55–56 (1968).
15 de Petris, S., Raff, M. C., and Maleucci, L., *Nature new Biol.*, **244**, 275–278 (1973).
16 Branton, D., *et al.*, *Science*, **190**, 54–56 (1975).
17 Falk, G., and Fatt, P., *J. Ultrastruct. Res.*, **28**, 41–60 (1969).
18 Cohen, A. I., *J. Cell Biol.*, **48**, 547–565 (1971).
19 Renthal, R., Steinemann, A., and Stryer, ..., *Expl Eye Res.*, **17**, 511–515 (1973).
20 Steinemann, A., and Stryer, L., *Biochemistry*, **12**, 1499–1502 (1973).
21 Romhanyi, G., and Molnar, L., *Nature*, **249**, 486–487 (1974).
22 Saari, J. C., *J. Cell Biol.*, **63**, 480–491 (1974).
23 Trayhurn, P., Mandel, P., and Virmaux, N., *Expl Eye Res.*, **19**, 259–265 (1974).
24 Yariv, J., Kalb, A. J., and Giberman, E., *J. molec. Piol.*, **85**, 183–186 (1974).

Role of nucleotide hydrolysis in microtubule assembly

MICROTUBULE assembly *in vitro* requires (in normal conditions) that GTP be bound to the exchangeable nucleotide-binding site of tubulin[1-3]. The bound GTP is hydrolysed during polymerisation and the resulting GDP remains bound to the microtubule while the phosphate is released into the medium. Although hydrolysis normally occurs during polymerisation, microtubules will form in a non-hydrolysable analogue of GTP, guanylyl imidodiphosphate (GMP-PNP)[4-6]. We have observed[6], as also has Arai[5], that microtubules polymerised in GMP-PNP are not depolymerised by concentrations of Ca which completely depolymerise GTP microtubules. We have now found that GMP-PNP microtubules do not display evidence for a rapid equilibrium between polymer and subunit. The irreversible behaviour of GMP-PNP microtubules may explain the insensitivity to Ca and suggests a function for nucleotide hydrolysis during tubulin polymerisation.

The reversibility of polymerisation was studied by dilution of preformed microtubules while constant solvent conditions were maintained. As others have observed[7-9], a plot of the extent of polymerisation in GTP against total tubulin concentration intercepts the concentration axis at about 0.2 mg ml[-1] (Fig. 1). This concentration has been referred to as the "critical concentration" and is the minimum concentration at which the polymer (microtubules) exists. According to the theory of Osawa and Kassai[10], the critical concentration is the point at which the first turn of the polymer helix is stable and can act as a nucleation structure for further growth. The mechanism of microtubule assembly is not known, but it is likely to be much more complicated than the simple condensation mechanism of Osawa and Kasai's theory. Thus the use of the term "critical concentration" does not imply here any specific model of polymerisation. The observation of a non-zero critical concentration, however, is important because it indicates that below this concentration no microtubules exist. In the dilution experiments this can only mean that the microtubules break down

Fig. 1 Dilution series of GTP and GMP-PNP microtubules. Microtubes were assembled at 35 °C in 0.1 M MES buffer (morpholino ethanesulphonic acid) at *p*H 6.6, 1 mM EGTA, 0.5 mM MgCl$_2$ and either 1.0 mM GTP or 1.5 mM GMP-PNP (concentrations at which maximum velocity of polymerisation is observed without inhibition). After an apparent plateau level of polymerisation had been achieved, the sample was diluted with the same buffer and the new turbidity level was determined after 5 min. Tubulin was prepared by three cycles of polymerisation and centrifugation in 25% glycerol. After the last centrifugation the pellet was resuspended in buffer and passed through Sephadex G-25 to remove most remaining traces of glycerol and nucleotide. ○, GTP microtubules; ●, GMP-PNP microtubules.

as the protein concentration is lowered, and therefore that they are in reversible equilibrium with subunits (or small aggregates). The complete breakdown of GPT microtubules after dilution has been confirmed by electron microscopy.

When GMP-PNP microtubules are subjected to the same dilution experiment, the data points lie on a straight line with an intercept on the concentration axis which is indistinguishable from zero (Fig. 1). In eight experiments at temperatures between 20 and 37 °C the average intercept was −0.07 mg ml[-1] and in no experiment was the intercept greater than zero. The failure of GMP-PNP microtubules to break down at very low concentrations up to 1 h after dilution has been confirmed by electron microscopy. At temperatures below about 18 °C GMP-PNP microtubules become reversible and Ca sensitive and the concentration curve has a positive intercept. At these lower temperatures the critical concentration of analogue tubules is always less than that of GTP tubules.

Although microtubules formed in GMP-PNP seem to be irreversible, or nearly so, complete polymerisation of tubulin is not observed in the analogue. The "plateau" level of polymerisation in GMP-PNP was always less than in GTP, although the difference varied widely between preparations. The lower level observed in the analogue is probably a result of inhibition of polymerisation by residual amounts of GDP remaining in the preparation. It is also possible that only a fraction of tubulin subunits are capable of polymerisation in GMP-PNP. Neither of these explanations alters the conclusions reached about the properties of GMP-PNP microtubules.

The polymerisation of tubulin and of muscle actin (and most cytoplasmic actins) requires bound nucleotide (ATP or ADP in the case of actin). In both proteins the bound nucleotide triphosphate is hydrolysed during polymerisation, and in both systems the non-hydrolysable imidodiphosphate analogue promotes polymerisation[11]. The role of bound nucleotide in actin polymerisation has been studied extensively but no clear concensus exists concerning the function of hydrolysis in polymerisation. It has been suggested that hydrolysis is involved in the depolymerisation reaction of actin[12] but the most direct test of this idea, depolymerisation of AMP-PNP F-actin by dilution, has not been performed successfully[11].

In the experiments reported here we have demonstrated that microtubules assembled in GMP-PNP are essentially irreversible and that hydrolysis of bound GTP probably promotes the depolymerisation of microtubules. By analogy, it is reasonable to assume that ATP hydrolysis has a similar function in actin polymerisation. The possible benefit of this to the living cell in regulating polymerisation of fibrous structures is apparent. By using the energy of binding of nucleotide triphosphate, rapid assembly can be obtained. Rapid polymerisation, however, may be incompatible with the need to disassemble rapidly the polymer later (such as during mitosis). Hydrolysis is thus used to produce a more readily depolymerisable state. It has been observed clearly in the case of microtubules and is probably also true of actin fibres that various degrees of stability of the polymer exist in the cell. At least some of this variation could be achieved by maintaining the bound nucleotide as the triphosphate. This could be done by the action of a nucleotide diphosphokinase ("transphosphorylase") such as has been observed as a component of tubulin preparations[2,6]. Future observations on the state of bound nucleotide in stable compared with labile microtubules and microfilaments may determine whether these speculations are valid.

This work was supported by a grant from the NSF to R.C.W.

R. C. WEISENBERG
W. J. DEERY

Department of Biology,
Temple University,
Philadelphia, Pennsylvania 19122

1 Weisenberg, R. C., *Science*, **177**, 1104–1105 (1972).
2 Jacobs, M., Smith, H., and Taylor, E. W., *J. molec. Biol.*, **89**, 455–468 (1974).
3 Kobayashi, T., *J. Biochem.*, **77**, 1193–1197 (1975).
4 Lockwood, A. H., Penningroth, S. M., and Kirschner, M. W., *Fedn. Proc.*, **34**, 540 (1975).
5 Arai, T., and Kaziro, Y., *Biochem. biophys. Res. Commun.*, **69**, 369–376 (1976).

* Weisenberg, R C , Deery, W J , and Dickinson, P , *Biochemistry* (in the press)
7 Gaskin, F , Cantor C R , and Shelanski, M L *J molec Brol* , **89**, 737–738 (1976)
8 Ohmsted, J B , Marcum, J M Johnson K A , and Borisy, G G , *J supramolec Struct* , 2, 429–450 (1974)
9 Engelborghs, Y , Heremans, K A H , DeMaeyer, L C M , and Hoobeka, J , *Nature* 259, 686–688 (1976)
10 Oosawa, F , and Kasai, M , *J molec Biol* , 4, 10–21 (1962)
11 Cooke, R , *Biochemistry* 13, 3250–3256 (1975)
13 Bender, N , Finkold, H , and Rack, M , *FEBS Lett* , 44, 209–212 (1974)

Protein evolution in cyanobacteria

THERE has been a great deal written in recent years on evolutionary relationships between prokaryotes and eukaryotes[1,2], and the possible origin of photosynthesis in eukaryotic cells by endosymbiosis of a cyanobacterium[2,3]. Molecular methods have been used in an attempt to elucidate the principal events in Precambrian cellular evolution. For example primary structures of plastocyanin[4], cytochrome f[5], and ferredoxin[6,7] have been published. These sequences have been compared to primary structures of functionally analogous macromolecules from eukaryotes. Before conclusions on the evolutionary relationships between prokaryotic cyanobacteria and eukaryotic algae and plants can be made from one representative amino acid sequence, it is necessary to evaluate the amount of amino acid sequence variation in proteins isolated from a wide range of cyanobacteria. In this study the amount of variation in plastocyanins and in cytochromes f was investigated. The results from complete and partial amino acid sequence determinations on these proteins from a number of cyanobacteria, suggest that the rate of evolution of the proteins in oxygenic photosynthetic prokaryotes is much less than the rates of evolution of corresponding proteins in eukaryotic algae and higher plants.

Plastocyanin and cytochrome f are components of the photosynthetic electron transfer chain[8]. Cyanobacteria were chosen which were reasonably diverse but also suitable to culture in the laboratory. The cytochromes f were from *Anabaena variabilis*, Kützing; *Plectonema boryanum* 1462/2 (also called *Phormidium luridum* var. olivaceum) and *Synechococcus* 6312. The first two are filamentous cyanobacteria and the third is unicellular. The nucleotide base compositions of these strains are 44% (ref. 9), 47% (ref. 9), and 50.2% (ref. 10) G + C respectively. The G + C content of *Spirulina maxima* was 36.6% (M. Herdmann, personal communication) measured on a sample of dried cells from the source in ref. 5. Plastocyanin from *P. boryanum* was also investigated. *A. variabilis* and *P. boryanum* were generally cultured in modified medium C of Kratz and Myers as described previously[4]. *P. boryanum* was also grown in the medium of Allen and Arnon[11] and latterly in the medium of Cannon *et al.*[12]. Cells of *Synechococcus* 6312 were a gift from R. Y. Stanier.

Isolation of the proteins was, in the initial stages, as described previously[4,13,14]. After ammonium sulphate fractionation of the cell extract and desalting to 1 mM phosphate buffer, pH 7, the solutions were passed down DEAE-cellulose where only *Synechococcus* 6312 cytochrome f was absorbed. Cytochromes f and plastocyanins from the other strains were absorbed on to CM-cellulose in the same buffer. The proteins were eluted from the ion-exchange celluloses with buffers of increased ionic strength. The proteins were further purified by absorption on to CM-cellulose at pH 3.9 followed by elution at increased pH and by gel filtration on Sephadex G-75. The protein preparations were checked for purity by polyacrylamide gel electrophoresis[15]. Later *Synechococcus* 6312 cytochrome f and *A. variabilis* plastocyanin were run in 8% sodium dodecyl sulphate (SDS) gels[17]. *A. variabilis* plastocyanin ran partly as a dimer after incubation with SDS in mild conditions although it ran as a monomer during Sephadex gel filtration with phosphate buffers. This phenomenon has been reported for spinach and *Scendesmus plastocyanin*[17].

Cytochrome f has been isolated from all cyanobacteria investigated so far in this laboratory but the preparation of plastocyanin has proved much more difficult. No plastocyanin was detected in *Synechococcus* 6312 including one preparation using 500 g fresh weight of cells. Attempts to prepare plas-

tocyanin on a small scale from another two strains of *Synechococcus*, 6307 and *Anacystis nidulans* (*Synechococcus* 6301)[10] and two strains of *Aphanocapsa* (unicellular cyanobacteria) were unsuccessful. A number of attempts to isolate plastocyanin on a large scale from *S. maxima* were also unsuccessful. Initial results of an immunological survey of the culture collection of R.Y. Stanier at Institut Pasteur (my unpublished results) have indicated the presence of plastocyanin in crude extracts of all heterocyst forming cyanobacteria investigated. Extracts from three unicellular species (2 *Aphanocapsa* and 1 *Microcystis*) and a few *Phormidium* and *Pseudoanabaena* species show definite cross reaction but tests on a large number of non-heterocyst forming filamentous cyanobacteria (including a species of *Spirulina*) and unicellular species proved negative. There was no immunological cross reaction with an extract of *A. nidulans* although the presence of plastocyanin in this strain has been suggested on the basis of electron paramagnetic resonance signals[16].

The amounts of the proteins isolated were generally very small for sequencing purposes (normal yields were a few mg from 1 kg fresh weight of cells). The sequence information was obtained from proteins isolated from cultures grown over a two-year period. In addition the gifts of some *Synechococcus* 6312 cytochrome f from G. Cohen-Bazire and *A. variabilis* cytochrome f from D. W. Krogmann were received.

P. boryanum cytochrome f was prepared in considerably higher yield (about 20 mg from 650 g fresh weight cells) from a batch grown in the medium of Cannon *et al.*[12], which contains 10–50 times more iron per l than most cyanobacterial media. This success could not be repeated. In an attempt to increase the yield of cytochrome f, batches of *A. variabilis* were grown in the presence of diphenylamine (12 mg l⁻¹) in modified Detmer's medium[19]. The strain did not grow well in these conditions and much lower yields of cells resulted, although the proportion of cytochrome f was increased. The copper content of the cyanobacterial plastocyanins, estimated with bathocuproine[20] indicated the presence of 1 mol Cu per mol protein.

Native oxidised plastocyanin and native reduced cytochrome f were degraded in a Beckmann, model 890 A automatic sequenator by a 70 min 'Quadrol' double cleavage programme[21]. All three cytochromes f tended to wash out of the cup and yields of phenylisothiocyanate derivatives dropped quickly. Identifications proved impossible after 20 cycles. Cytochromes f from *Synechococcus* 6312 and *P. boryanum* were later degraded using a single cleavage programme with dimethylbenzylamine[22]. Identification of phenylisothiocyanate derivatives was carried out as described previously[4] with the addition that some derivatives were hydrolysed to the free amino acids with HCl[23]. The criteria of reliability of sequenator results adopted were generally those of Niall[24].

Some of the residues and amide assignments at the amino

	P. boryanum	Spirulina maxima	Monochrysis lutheri	Porphyra tenera	Euglena gracilis	Alaria esculenta
P. boryanum	—	56	46	54	43	47
Spirulina maxima	56	—	47	53	42	49
Monochrysis lutheri	46	47	—	48	35	45
Porphyra tenera	54	53	48	—	35	67
Euglena gracilis	43	42	35	35	—	39
Alaria esculenta	47	49	45	67	39	—

For the comparison Asx is taken as identical to both Asp and Asn, and Glx to both Glu and Gln.

• terminal regions of the cytochromes *f* were confirmed by digestion of small amounts of the cytochromes with trypsin after removal of the haem. The rest of the available material (about 8 mg) was used. Peptides were isolated by established methods for prokaryotic type *c* cytochromes[25]. Yields of tryptic peptides from many parts of the sequences were very low or zero partly due to the presence of –Asn–Gly– sequences which readily undergo deamidation[25].

The alignment of the amino-terminal regions of the cytochromes *f* is shown in Fig. 1. The other parts of the sequences from which peptides were obtained and sequenced are aligned by homology with the complete sequences of *P. boryanum* cytochrome *f* (my unpublished results); *Porphyra tenera* cytochrome *f*[26] and the other prokaryotic and eukaryotic cytochromes *f* aligned previously[4].

The similarity matrix (Table 1) shows that *P. boryanum* and *Spirulina maxima* cytochromes *f* have nearly identical percentages of similarity to each eukaryotic algal sequence taken in turn. The cyanobacterial sequences are 10% closer to each other than the average for all the other sequence comparisons. They are much closer to the red algal (*P. tenera*) cytochrome *f*[26] than to the other eukaryotic algal cytochromes. This is interesting in view of the many close similarities between cyanobacteria and red algae[27]. The partial sequences of the other cytochromes *f* strongly suggest that these numerical relationships are general.

The amino terminal sequence of *P. boryanum* plastocyanin is aligned in Fig. 2 with those of *A.variabilis* plastocyanin[4], *Chlorella fusca* plastocyanin[18] and two examples of higher plant plastocyanins[29,30]. The partial sequence of plastocyanin

from *P. boryanum* has considerable similarity to the same regions in the other plastocyanins if a deletion at position 13 is considered. Gly-34 (one of the tentative assignments), Val-33 and Met-7 have previously been invariant residues. The similarity matrix (Table 2) gives percentage comparisons between the amino-terminal sequences. This region of *A.variabilis* plastocyanin is slightly closer to the eukaryotic algal plastocyanin than to *P. boryanum* plastocyanin. This finding is valid even if the changes in the previously invariant residues are treated with caution.

Amino acid sequences of cyanobacterial ferredoxin from *Aphanothece* and two strains of *Spirulina*[6,7] have been compared to primary structures of ferredoxins from eukaryotic algae and plants. Results were obtained similar to those in the present study. The prokaryotic sequences are either not closer or only slightly closer to each other than to the corresponding eukaryotic algal sequences. Where comparisons with higher plant sequences can be made (in the cases of plastocyanin and ferredoxin) cyanobacterial primary structures do seem to be closer to each other than to the higher plant sequences. Ferredoxin sequence comparisons ought to be treated with caution because multiple ferredoxins with quite distinct amino acid compositions have been shown to be present in some plants[8] and in cyanobacteria[31,32].

Fossil evidence indicates that filamentous and unicellular forms of cyanobacteria diverged in the early Precambrian era, about 3×10^9 yr ago[33] and many of these have subsequently evolved little. In the Gunflint Iron formations (1.6×10^9–2×10^9 yr old) many representatives of extant cyanobacterial families (Chroococcaceae, Oscillatoriaceae and Nostocaceae)

Fig. 1 Alignment of cytochromes *f. a, Synechococcus* 6312; *b, Anabaena variabilis; c, Plectonema boryanum; d, Spirulina maxima; e, Monochrysis lutheri* (chrysophytan alga); *f, Porphyra tenera* (red alga); *g, Euglena gracilis; h, Alaria esculenta* (brown alga).

	1	10	20	30
a.	Ala-Asp-Ile-Ala-Asp-Gly-Ala-Lys-Val-Phe-Ser-Ala-Asn-Cys-Ala-Ala-Cys-His-Met-Gly-Gly-Gly-Asn(Val)		(Val)Met-Ala-Asn-Lys-Thr-Leu-	
b	Ala-Asp-Ser-Val-Asn-Gly-Ala-Lys-Ile-Phe-Ser-Ala-Asn-Cys-Ala-Ser-Cys-His-Ala-Gly-Gly-Lys		Thr-Leu-	
c	Ala-Asp-Ala-Ala-Ala-Gly-Gly-Lys-Val-Phe-Asn-Ala-Asn-Cys-Ala-Ala-Cys-His-Ala-Ser-Gly-Gly-Gly-Gln-	Ile-Asn-Gly-Ala-Lys-Thr-Leu-	
d	Gly-Asp-Val-Ala-Ala-Gly-Ala-Ser-Val-Phe-Ser-Ala-Asn-Cys-Ala-Ala-Cys-His-Met-Gly-Gly-Arg-Asn-Val-_Ile-Val-Ala-Asn-Lys_Thr-Leu-			
e	Gly-Asp-Ile-Ala-Asn-Gly-Glu-Gln-Val-Phe-Thr-Gly-Asn-Cys-Ala-Ala-Cys-His-Ser-Val-Glx-Glx-Glx-Val-	Thr-Leu-Glu-Leu-Ser-Ser-Leu-	
f	Ala-Asp-Leu-Asp-Asn-Gly-Glu-Lys-Val-Phe-Ser-Ala-Asn-Cys-Ala-Ala-Cys-His-Ala-Gly-Gly-Asn-Asn-Ala-	Ile-Met-Pro-Asp-Lys-Thr-Leu-	
g	Gly-Gly-Ala-Asp-Val-Phe-Ala-Asp-Asn-Cys-Ser-Thr-Cys-His-Val-Asn-Gly-Gly-Asn-Val-	Ile-Ser-Ala-Gly-Lys-Val-Leu-	
h	Ile-Asp-Ile-Asn-Asn-Gly-Glu-Asn-Ile-Phe-Thr-Ala-Asn-Cys-Ser-Ala-Cys-His-Ala-Gly-Gly-Asn-Asn-Val-	Ile-Met-Pro-Glu-Lys-Thr-Leu-	

	40	50	60
a			Asn-Ala-Met-
b	-Lys-Lys-Ala-Asp-Leu-Glu-Lys-Tyr		Asn-Ala-Met-
c	-Lys-Lys-Asn-Ala-Leu-Thr-Ala————————Asn-Gly-Lys-Asp-Thr-Val-Glu-Ala-Ile-Val-Ala———Gln-Val-	Thr-Asn-Gly-Lys-Gly-Ala-Met-
d	-Ser-Lys-Ser-Asp-Leu-Ala-Lys-Tyr-Leu-Lys-Gly-Phe-Asp-Asp-Asp-Ala-Val-Ala-Ala-Val-Ala-Tyr-Gln-Val-	Thr-Asn-Gly-Lys-Asn-Ala-Met-
e	-Trp-Lys————————Ala-Lys-Ser-Tyr-Leu-Ala-Asn-Phe-Asn-Gly-Asp-Glu-Ser-Ala-Ile-Val————Tyr-Gln-Val-	Thr-Asn-Gly-Lys-Asn-Ala-Met-
f	-Lys-Lys————Asp-Val————————Leu-Glu-Ala-Asn-Ser-Met-Asn-Thr-Ile-Asp-Ala-Ile-Thr————Tyr-Gln-Val-	Gln-Asn-Gly-Lys-Asn-Ala-Met-
g	-Ser-Lys-Thr-Ala-Ile-Glu-Glu-Tyr-Leu-Asp-Gly-Gly-Tyr————Thr-Lys-Glu-Ala-Ile-Glu————Tyr-Gln-Val-	Arg-Asn-Gly-Lys-Gly-Pro-Met-
h	-Lys-Lys————Asp-Ala————————Leu-Ala-Asp-Asn-Lys-Met-Val-Ser-Val-Asn-Ala-Ile-Thr————Tyr-Gln-Val-	Thr-Asn-Gly-Lys-Asn-Ala-Met-

	70	80	90
a	-Pro-Gly-Phe-Ala-Gly-Arg		
b	-Pro-Ala-Phe-Lys-Gly-Arg-Leu-Lys-Pro-Glu-Glu-Ile-Glx-Asx-Val-Ala-Ala-Tyr-Val-Leu-Gly-Lys-Ala-Asp-	Ala-Asp-Trp-Lys
c	-Pro-Ala-Phe-Lys-Gly-Arg-Leu-Ser-Asp-Asp-Gln-Ile-Gln-Ser-Val-Ala-Leu-Tyr-Val-Leu-Asp-Lys-Ala-Glu-	Lys-Gly-Trp
d	-Pro-Gly-Phe-Asn-Gly-Arg-Leu-Ser-Pro-Lys-Gln-Ile-Glu-Asp-Val-Ala-Ala-Tyr-Val-Val-Asp-Gln-Ala-Glu-	Lys-Gly-Trp
e	-Pro-Ala-Phe-Gly-Gly-Arg-Leu-Glu-Asp-Asp-Glu-Ile-Ala-Asx-Val-Ala-Ser-Tyr-Val-Leu-Ser-Lys-Ala-Gly	
f	-Pro-Ala-Phe-Gly-Gly-Arg-Leu-Val-Asp-Glu-Asp-Ile-Glu-Asp-Ala-Ala-Asn-Tyr-Val-Leu-Ser-Gln-Ser-Glu-	Lys-Gly-Trp
g	-Pro-Ala-Trp-Glu-Gly-Val-Leu-Ser-Glu-Asp-Glu-Ile-Val-Ala-Val-Thr-Asp-Tyr-Val-Tyr-Thr-Gln-Ala-Gly-	Gly-Ala-Trp-Ala-Asn-Val
h	-Pro-Ala-Phe-Gly-Ser-Arg-Leu-Ala-Glu-Thr-Asp-Ile-Glu-Asp-Val-Ala-Asn-Phe-Val-Leu-Thr-Gln-Ser-Asp-	Lys-Gly-Trp-Asp

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |

a' Asp-Thr-Val-Lys-Val-Ile-Met-Gly-Gly-(Ser)-Lys-Gly————Leu-Val-Phe-Glu-Pro-Ala-Val-Val-Asn-Val-

b Glu-Thr-Tyr-Thr-Val-Lys-Leu-Gly-Ser-Asp-Lys-Gly-Leu-Leu-Val-Phe-Glu-Pro-Ala-Lys-Leu-Thr-Ile-

c Asp-Val-Thr-Val-Lys-Leu-Gly-Ala-Asp-Ser-Gly-Ala-Leu-Val-Phe-Glu-Pro-Ser-Ser-Val-Thr-Ile-

d Leu-Asp-Val-Leu-Leu-Gly-Gly-Asp-Asp-Gly-Ser-Leu-Ala-Phe-Ile-Pro-Gly-Asn-Phe-Ser-Val-

e Leu-Glu-Val-Leu-Leu-Gly-Ser-Gly-Asp-Gly-Ser-Leu-Val-Phe-Val-Pro-Ser-Glu-Phe-Ser-Val-

| | 24 | 25 | 26 | 27 | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | 37 | 38 | 39 | 40 | 41 | 42 | 43 | 44 |

a -Lys-Ala-Gly-Asp-Thr-Ile-Gln-Phe-Glu-Val-(Gly)-Gln-(Leu)-Pro-Pro-His-

b -Lys-Pro-Gly-Asp-Thr-Val-Glu-Phe-Leu-Asn-Asn-Lys-Val-Pro-Pro-His-Asn-Val-Val-Phe-Asp-

c -Lys-Ala-Gly-Glu-Thr-Val-Thr-Trp-Val-Asn-Asn-Ala-Gly-Phe-Pro-His-Asn-Ile-Val-Phe-Asp-

d -Ser-Ala-Gly-Glu-Lys-Ile-Thr-Phe-Lys-Asn-Asn-Ala-Gly-Phe-Pro-His-Asn-Val-Val-Phe-Asp-

e -Pro-Ser-Gly-Glu-Lys-Ile-Val-Phe-Lys-Asn-Asn-Ala-Gly-Phe-Pro-His-Asn-Val-Val-Phe-Asp-

Fig. 2 N-terminal amino acid sequences of plastocyanins. *a*, *Plectonema boryanum* 1462/2; *b*, *Anabaena variabilis* (Kützing); *c*, *Chlorella fusca* (vacuolata); *d*, Potato (*Solanum tuberosum*); *e*, French bean (*Phaseolus vulgaris*). Residues shown in brackets are tentative.

have been identified[22]. In more recent rocks (the Bitter Spring formations, 0.9×10^9 yr old) many species of cyanobacteria have been identified at the specific or generic level with extant organisms, including *Spirulina*. If the strains of cyanobacteria for which protein sequence information is available are as evolutionarily diverse as the fossil record suggests, then it is perhaps surprising that their amino acid sequences are not more different. Cyanobacteria were a diverse group of organisms 3×10^9 yr ago[33] and eukaryotic algae diverged in a time measured in hundreds of millions of years[33]. There seems to be, however, the same degree of similarity in analogously functioning proteins. It has been postulated that the rate of evolution of a particular protein in a wide range of organisms is approximately constant[34]. This has been shown for cytochrome *c*, haemoglobin and fibrinopeptide and has enabled phylogenetic trees to be constructed which correspond well with palaeontological evidence[34]. If cyanobacteria have evolved very little[33] in morphology and physiology since the oxygenic mode of photosynthesis was established in the Precambrian era there is probably little selective value for changes in primary structure of their photosynthetic proteins. Chloroplasts have evolved and diverged in a wide variety of eukaryotes over a much shorter time span than that in which cyanobacteria have

continued to populate similar ecological habitats. The rate of evolution of cytochrome *f* (a type *c* cytochrome) could be expected to be comparable to that of mitochondrial cytochrome *c* which is three "accepted point mutations" per 10^8 yr (ref. 34). The rate of evolution of higher plant plastocyanin is about twice this value[35]. At these rates of change of amino acid sequence one would expect very little remaining similarity in primary structure of these proteins from diverse blue-green algae (except for a few highly invariant amino acid residues required for a specific function such as binding of the prosthetic group). When the cyanobacterial proteins are compared with analogously functioning proteins in other cyanobacteria or in eukaryotes it would seem that their rates of evolution are much less than the rates of evolution of the proteins in eukaryotic algae and higher plants. The results seem very much against the hypothesis of neutral mutation in protein[36]. This is the hypothesis that most evolutionary change in DNA is not due to Darwinian natural selection but to selectively neutral mutations which become passively fixed through the action of random genetic drift. These selectively neutral mutations in DNA result in substitutions of amino acids in proteins which do not cause appreciable changes in protein function. In fact this does not seem to be the case here. It would seem that the role of neutral mutation is small and the amount of change in these cyanobacterial amino acid sequences which has occurred seems to reflect the functional differences between the proteins in the variety of organisms. For example it was seen that the cyanobacterial and red algal cytochrome *f* were particularly close.

It is also possible that the pattern of divergence in cyanobacterial primary structures has been obscured by transfer of genetic information coding for these photosynthetic proteins. There is evidence to suggest that this has occurred between distantly related bacteria[37], but whatever the reason for the sequence comparisons obtained, protein sequence studies have produced a considerable amount of evidence for the hypothesis of the common origin of oxygenic photosynthesis in prokaryotes and eukaryotes. It seems unlikely, however, that molecular methods will enable construction of phylogenetic trees which would indicate the precise manner in which cyanobacteria evolved into chloroplasts or chloroplast-containing eukaryotes.

I thank Dr R. P. Ambler for invaluable advice and assistance

Table 2 Similarity matrix for plastocyanins

	P. boryanum	*A. variabilis*	*Chlorella fusca*	Potato	French bean
P. boryanum	—	49	41	36	33
A. variabilis	49	—	56	36	38
Chlorella fusca	41	56	—	49	46
Potato	36	36	49	—	69
French bean	33	38	46	69	—

The values are expressed as a percentage identity for residues 1–39. The gap in the *Plectonema* sequence has been assigned as a 21st amino acid.

• and Mrs M. Daniel for performing sequenator runs. I thank the MRC for a scholarship and the following for gifts of material: Dr G. Cohen-Bazire and Professor D. W. Krogmann for cytochromes *f*; Professor R. Y. Stanier for cells of *Synechococcus* 6312 and Dr N. G. Carr for the strain of *A. variabilis*.

ALASTAIR AITKEN*

Department of Molecular Biology,
University of Edinburgh,
Edinburgh EH9 3JR, UK

1 Allsopp, A., *New Phytol.*, 68, 591–612 (1969)
2 Margulis, L., *Origin of Eukaryotic Cells* (Yale University Press, New Haven, 1970)
3 Taylor, F. J. R , *Taxonomy*, 23, 229–258 (1974)
4 Aitken, A., *Biochem. J.*, 149, 675–683 (1975).
5 Ambler, R. P , and Bartsch, R G., *Nature*, 253, 285–288 (1975).
6 Tanaka, M., Haniu, M , Yasunobu, K T., Rao, K. K., and Hall, D. O., *Biochemistry*, 14, 5535–5540 (1975).
7 Wada, K., Hase, T., Tokunaga, H , and Matsubara, H., *FEBS Lett.*, 55, 102–104 (1975).
8 Bishop, N. L., *A. Rev. Biochem* , 40, 197–226 (1971).
9 Edelman, M , Swinton, D , Schiff, J. A., Epstein, H T , and Zeldin, B., *Bact. Rev.*, 31, 315–331 (1967).
10 Stanier, R Y., Kunisawa, R , Mandel, M., and Cohen-Bazire, G., *Bact Rev.*, 35, 171–205 (1971)
11 Allen, M. B., and Arnon, D. I., *Pl Physiol , Lancaster*, 30, 366–372 (1955).
12 Cannon, R E., Shane, M S , and Bush, V N., *Virology*, 45, 149–153 (1971).
13 Lightbody, J. J., and Krogmann, D W , *Biochim. biophys. Acta*, 131, 508–515 (1967).
14 Biggins, J., *Pl Physiol., Lancaster*, 42, 1447–1456 (1967).
15 Ornstein, L., *Ann N Y. Acad. Sci* , 121, 321–349 (1964).
16 Weber, K., and Osborn, M., *J. biol Chem* , 244, 4406–4412 (1969).
17 Stegalman, M H , Rashed, I , and Boger, P , *Biochim. biophys. Res. Commun.*, 65, 1456–1463 (1975).
18 Visser, J. W. M., Amesz, J , and Van Gelder, B F , *Biochim. biophys. Acta*, 333, 279–287 (1974).
19 Ogawa, T., and Vernon, L P , *Biochim biophys. Acta*, 226, 88–97 (1971).
20 Wharton, D. C , and Rader, M , *Analyt. Biochem.*, 33, 226–229 (1970).
21 Edman, P., and Begg, G , *Eur. J Biochem* , 1, 80–91 (1967)
22 Hermodson, M A , Ericsson, L H , Titani, K , Neurath, H , and Walsh, K A , *Biochemistry*, 11, 4493–4502 (1972).
23 Van Orden, H. O., and Carpenter, F. M., *Biochem biophys Res. Commun* , 14, 399–403 (1964).
24 Niall, H. D., *Meth. Enzymol* , 27, 942–1010 (1973)
25 Ambler, R. P , *Biochem J* , 89, 349–378 (1963)
26 Laycock, M. V., *Biochem J* , 149, 271–279 (1975)
27 Wolk, C. P , *Bact. Rev* , 37, 32–101 (1973)
28 Kelly, J., and Ambler, R. P , *Biochem. J* , 143, 681–690 (1974).
29 Ramshaw, J. A. M , Scawen, M. D , Bailey, C J , and Boulter, D., *Biochem J* , 139, 583–592 (1974)
30 Milne, P. R., Wells, J. R. E , and Ambler, R P., *Biochem. J.*, 143, 691–701 (1974)
31 Hutson, K G., and Rogers, L. J., *Biochem. Soc. Trans* , 3, 377–379 (1975).
32 Hase, T., Wada, K , and Matsubara, H , *J. Biochem. (Tokyo)*, 78, 605–610 (1975)
33 Schopf, J W., *Origin of Life*, 5, 119–135 (1974)
34 Dayhoff, M. O., *Atlas of Protein Sequence and Structure*, 4 (National Biomedical Research Foundation, Washington, DC, 1972)
35 Peacock, D., and Boulter, D., *J molec Biol* , 95, 513–527 (1975)
36 King, J. L., and Jukes, T. H., *Science*, 164, 788–798 (1969).
37 Ambler, R. P , *Syst Zool.*, 22, 554–565 (1973)

New approach to steroid conversion using activated immobilised microorganisms

THE marked increase in demand for contraceptives and anti-inflammatory agents such as cortisol and prednisolone, combined with a diminishing supply of steroid raw materials may lead to shortages of steroid drugs[1]. Thus it is important to develop new sources for steroids as well as to devise more efficient means for steroid conversions. Here we report a new approach to steroid transformation in which activated immobilised microorganisms are utilised and which represents a promising alternative to conventional microbial transformation processes.

Immobilised microorganisms have attracted increasing interest as catalysts in the past few years[2–4] They have the same operational advantages as those inherent in immobilised enzymes[5], they are reusable, they are well suited for continuous operation in controlled conditions and further, there is no stringent demand for asepsis during opera-

tion. Immobilised microorganisms offer the additional advantage that tedious and costly enzyme isolation is obviated, that the enzyme is usually more stable due to its localisation in its "natural environment", and that there is usually no need for cofactor. Thus, provided competing reactions can be eliminated, immobilised microorganisms are very promising catalysts.

In this communication we consider especially a property unique to immobilised whole-cell catalysts—the possibility of *in situ* activation of the immobilised enzyme activity. The system studied was the important corticosteroid transformation cortisol $\xrightarrow{\Delta^1\text{-dehydrogenase}}$ prednisolone. The reaction was catalysed by polyacrylamide-entrapped *Corynebacterium simplex*.

C. simplex cells were grown in a medium of 0.25% yeast extract; the Δ^1-dehydrogenase activity being induced by addition of cortisol to the culture 12 h before collection by continuous centrifugation at 10,000g. The cells (5 g wet weight) were suspended in 20 ml of ice-cold 0.1 M Tris-HCl buffer, pH 7.5, and mixed with 25 ml of ice-cold aqueous monomer solution containing 7.13 g of acrylamide and 0.38 g of *N,N'*-methylene-*bis*-acrylamide. The mixture was poured into a sandwich-like polymerisation vessel (made of two glass plates 20×20×0.2 cm, spaced 2 mm apart with a piece of latex tubing) and the catalyst potassium persulphate (50 mg) and tetramethylenediamine (100 mg) were added in water (1 ml). Nitrogen gas was bubbled through the suspension and polymerisation started within 2 min. The polyacrylamide gel sheet was fragmented in a blender and the gel granules (average size 0.2 mm) were washed extensively with Tris buffer and then stored at −20 °C at which temperature the preparation was stable for several months.

The 3-ketosteroid-Δ^1-dehydrogenase activity of the immobilised *C. simplex* was assayed conveniently spectrophotometrically (Fig 1), and sole product, prednisolone, was identified by thin-layer chromatography. Approximately 40% of the dehydrogenase activity was retained during the immobilisation procedure (all bacteria added were immobilised and no release of bacteria was observed during incubations). Initial experiments revealed, however, that the activity declined rather rapidly on repeated batchwise conversions of high loads of cortisol and this could only be compensated for to a limited degree by addition of the artificial electron acceptor menadione[6]. Instead the stabilising influence of various nutrients and salts was investigated; the results are given in Table 1. In media

Fig. 1 Assay of Δ^1-dehydrogenase activity. Δ^1-Dehydrogenase activity of immobilised *C. simplex* was determined in air-saturated 1 mM cortisol. The progress of the reaction was monitored at 285 nm by pumping the supernatant through a flow cell ($\Delta\varepsilon = 2,820$ M^{-1} cm^{-1}) The initial transformation rates listed in Table 1 were determined in an assay composition of 0.5 g of gel granules (wet weight)+9.0 ml of 0.05 M Tris-HCl, pH 7.0, plus, 0.5 ml of 20 mM cortisol (methanol).
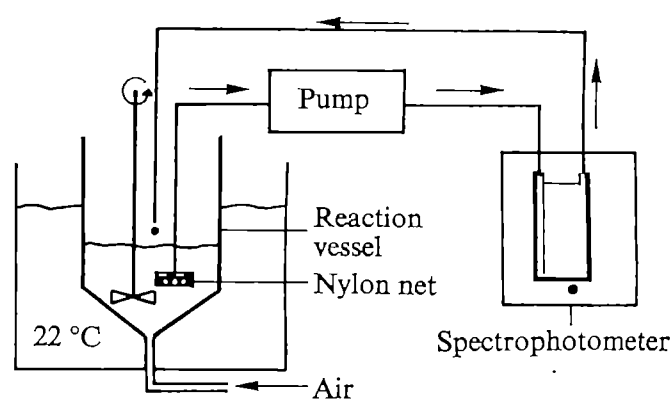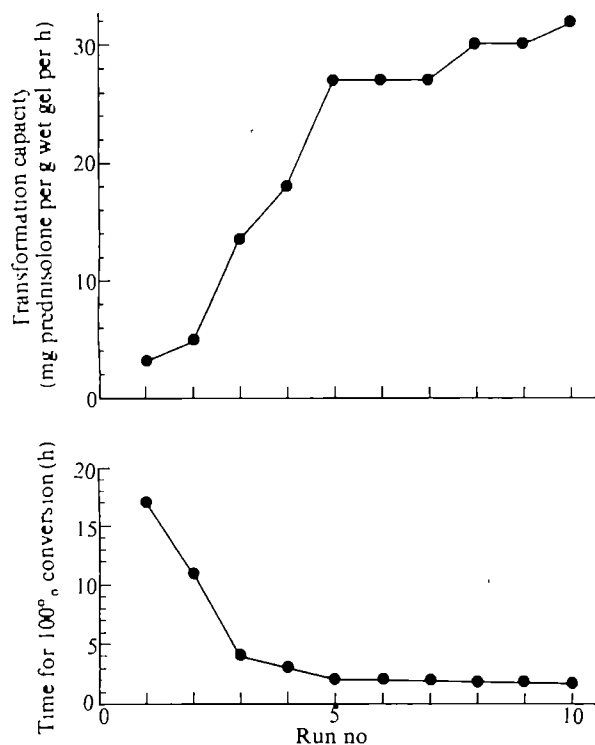
Fig. 2 Repeated batchwise transformation of cortisol to prednisolone. *C. simplex* gel (2.0 g, wet weight) was suspended in 285 ml of 0 5% peptone, pH 7.0, +15 ml of 20 mM cortisol (methanol). The transformation was carried out in the apparatus depicted in Fig. 1. When 100% conversion to prednisolone was reached (indicated by the location of the points), the gel was washed and again incubated with fresh cortisol-containing medium. The whole experiment lasted 4 d.

consisting of water or buffer the activity decreased, whereas in peptone-containing and glucose-containing media the activity was not only preserved but also increased to several times that of the original activity.

The 0.5% peptone medium and the 0 1% peptone +0.2% glucose medium were selected for further study and an experiment with repeated batchwise transformation was conducted. Both media were approximately equally efficient, and in Fig. 2 the results obtained with the 0.5% peptone medium are depicted. As can be seen, the transformation capacity increased remarkably with each run, thus while in the first batch 100% transformation was obtained after 18 h, the last batch was completed in less than 2 h. The transformation capacity at the end of the experiment was approximately 0.5 g of steroid per d per g of gel (wet weight).

This transformation capacity means that a continuously operating reactor loaded with only 1–2 kg of immobilised *C. simplex* gel would suffice to supply the Swedish demand for 1-dehydrogenated steroids (8 million inhabitants consuming approximately 250 kg of 1-dehydrogenated steroids per year at a value of £5 million). The calculation is based on the assumptions that the catalyst is completely stable, that no decrease in efficiency occurs on scaling up and that the dehydrogenation is the final step in the production of Δ^1-dehydrogenated steroids. Even if these assumptions are too optimistic, the calculation still serves to give some idea of the very modest reactor dimensions needed even for large-scale production of Δ^1-dehydrognated steroids.

Preliminary experiments show that the so-called pseudo-crystallofermentation technique[7] is applicable also to entrapped *C simplex*. Cortisol was thus added in an amount (3.6 g l⁻¹) far exceeding its solubility in the medium and was

Table 1 Activating effect of nutrients on the Δ^1-dehydrogenase activity

Medium	Initial transformation rate (%) after				
	0	2	6	10	16 d
Peptone, 0.5%	100	460	500	650	530
Peptone, 0.1% + Glucose, 0.2%	100	270	230	250	250
Glucose, 0.2%	100	210	170	110	90
K_2HPO_4, 0.1 M, pH = 7.0	100	70	50	60	60
Tris-HCl, 0.05 M, pH = 7.0	100	100	70	70	30
H_2O	100	60	40	40	20

C. simplex gel (0 5 g) was incubated in 9.0 ml of medium as indicated and 0.5 ml of 20 mM cortisol (methanol) was added. The suspension was shaken on a rotary shaker at 25 °C and at ₄8-h intervals the medium was replaced by fresh cortisol-containing medium. At the times indicated the gel was filtered off, washed and assayed for Δ^1-dehydrogenase activity as described in Fig. 1.

completely converted at approximately the same rate as in experiments with dissolved cortisol. The product prednisolone, which precipitated out, could be isolated simply by filtration after the rather dense gel granules had been allowed to settle. This technique allows reduction of media volumes by orders of magnitude and thus also of nutrients.

The reason for the observed activation is not known, it could be microbial growth or protein synthesis Microscopic studies of gel granules revealed that there is at least some increase in the amount of entrapped bacteria. Continuous induction leading to *de novo* synthesis of Δ^1-dehydrogenase in existing cells is also a plausible explanation. Cell lysis, which might enhance substrate/product transport, is, on the other hand, considered a rather unlikely cause of activation, since increased activity was observed only in media containing nutrients (Table 1).

In situ (re)activation procedures similar to that given here should obviously also be applicable to other processes catalysed by immobilised microorganisms, for example, biotransformation of other steroids and antibiotics as well as other compounds.

P. O Larsson
S Ohlson
K. Mosbach

Biochemical Division,
Chemical Center, PO Box 740,
S-220 07 Lund 7, Sweden

[1] Applezweig, N., *Chem Week*, July 10, 31–36 (1974)
[2] Sato, T, *et al*, *Biotechnol Bioengng*, 17, 1797–1804 (1975)
[3] Sami, R, and Vieth, W R, *J appl Chem Biotechnol*, 25, 115–141 (1975).
[4] Mosbach, K., and Larsson, P O, *Biotechnol Bioengng*, 12, 19–27 (1970)
[5] Mosbach, K, *FEBS Lett*, 62, Suppl, E80–E95 (1976)
[6] Lee, B K, Brown, W. E, Ruy, D Y, and Thoma, R W., *Biotechnol Bioengng*, 13, 503–515 (1971)
[7] Kondo, F., and Masuo, E, *J gen appl. Microbiol*, 7, 113–117 (1961)

Polymers for the sustained release of proteins and other macromolecules

Since the first demonstration that silicone rubber[1] could be used as an implantable carrier for sustained delivery of low molecular weight compounds in animal tissues, various drug delivery systems have been developed. But except for the reports of Davis[2,3] and Gimbrone *et al.*[4], there has been little success in the development of slow release agents for large molecular weight compounds. Furthermore, the polymers used in those studies, polyvinylpyrrolidone and polyacrylamide, are often inflammatory in animal tissues and

Table 1 Host response to corneal polymer gel implants

Polymer	No. of tests	Inflammation*		
		None	Mild	Significant
Polyacrylamide, 20%	10	20%	30%	50%
Polyvinylpyrrolidone, 20%	4	0	0	100%
Polyvinylalcohol (unwashed, 1%, 5%, 10%)	20	15%	70%	15%
Polyvinylalcohol (washed, 10%)	6	67%	33%	0
Hydron-S (with and without water in the polymer casting solution)	15	100%	0	0
Ethylene–vinyl acetate copolymer (unwashed)	20	40%	60%	0
Ethylene–vinyl acetate copolymer (washed)	20	100%	0	0

*The criteria were oedema, white cells and neovascularisation as observed through a stereomicroscope. A response was judged to be mild if any one of the three characteristics was detected, and significant if their presence was enough to cause corneal opacity.

usually permit only brief periods of sustained release. We now present a simple method for incorporating various proteins and other macromolecules into non-inflammatory polymers. Sustained release of biochemically active macromolecules has been demonstrated for periods exceeding 100 d.

Sterile polymer casting solutions were made by dissolving polymers in appropriate solvents: Hydron (a polymer of hydroxyethylmethacrylate) in absolute alcohol at 37 °C; ethylene–vinyl acetate copolymer (40% by weight vinyl acetate) in methylene chloride at 37 °C; and polyvinylalcohol by autoclaving in distilled water. Slow release pellets were made by placing a mixture of casting solution and the macromolecule to be released in a conical mould, 2–4 mm in diameter and 1.5 mm deep. Glass moulds were used for the ethylene–vinyl acetate copolymer because the polymer casting solution reacted with plastics causing adhesion. The moulds were dried under a mild vacuum (600 mmHg) overnight causing the solvent to evaporate leaving the macromolecule trapped within the polymer matrix. To retard further release of macromolecules, 'sandwiches' were sometimes made by coating the matrix with pure polymer in casting solution. Coating was achieved by dipping the polymer pellet in a puddle of pure polymer in casting solution and drying. In an alternative method of coating, pure polymer in casting solution was dried in the bottom of a mould, a middle layer of polymer and macromolecule was added and dried, and finally a top layer of pure polymer in casting solution was added and dried. Before use, the dried slow-release pellets were rehydrated by adding a drop of Ringer solution.

To study tissue response to the polymers, sterile polymer pellets (1.5 × 1.5 × 0.5 mm³) were implanted in the eyes of rabbits, by aseptically creating a small pocket in the cornea⁴. Inflammation was judged by stereomicroscopy. The sensitivity of the cornea to inflammatory stimuli as well as its clarity and avascularity make it an excellent site to observe inflammatory characteristics: oedema, white-cell infiltration, and neovascularisation⁵.

Ethylene–vinyl acetate copolymer, exhaustively washed in alcohol, and Hydron caused no inflammation in the cornea (Table 1). Histological sections made 3–4 weeks after implantation showed no foreign cells in the region surrounding the polymers or in other areas of the cornea. In contrast, polyacrylamide and polyvinylpyrrolidone, prepared according to Davis⁶, were significantly inflammatory. Polyacrylamide has been used in our laboratory for several years⁴, and inflammatory responses have been reduced by various washing techniques (R. Arensman and M. Gimbrone, unpublished). Unwashed polyvinylalcohol pellets also caused inflammation. Washing the polymer exhaustively with alcohol greatly reduced, but did not eliminate, the inflammatory response (Table 1).

Polyvinylalcohol, Hydron and ethylene–vinyl acetate copolymer were examined for the ability to release soybean trypsin inhibitor (molecular weight 21,000). The polymer–water ratio in the casting solution of polyvinylalcohol was

important in the determination of release rates. The effects of various polyvinylalcohol concentrations in the casting solution on the rate of release of soybean trypsin inhibitor are summarised in Fig. 1. An increase in the polymer concentration in the casting solution significantly decreased the initial rate at which protein was released. A threshold level was apparently reached at about 10% polyvinylalcohol, because further increases in polymer concentration did little to retard release from the pellet. Water was also added to the Hydron–alcohol–protein casting solution. The greater the water–polymer ratio, the greater was the rate of protein release from the pellet.

Figure 1 illustrates two additional features of these slow release systems: (1) a large percentage of protein was released during the first hour of incubation, that is, 'burst effect'; (2) the rate of release was significantly retarded in polymer 'sandwiches'.

Figure 2a shows the release of soybean trypsin inhibitor from the three polymers. Protein was released from Hydron relatively rapidly, more slowly from polyvinylalcohol, and least rapidly from ethylene–vinyl acetate copolymer. The 'burst' was evident for all three systems.

Fig. 1 Release of soybean trypsin inhibitor from polyvinylalcohol. The percentage of polymer in the casting solution was varied. Protein concentration in all casting solutions was 12 mg ml⁻¹. Release characteristics were studied by incubating polymer pellets in small test tubes with lactated Ringer solution at 37 °C. Periodically, pellets were removed to new tubes containing fresh Ringer solution. Protein release from the pellets was assayed by the Lowry method⁸. Polyvinylalcohol, △, 1.2%; ○, 4.8%; ▲, 6%; ■, 10%; circle and cross, 20%; ●, 10% sandwich.
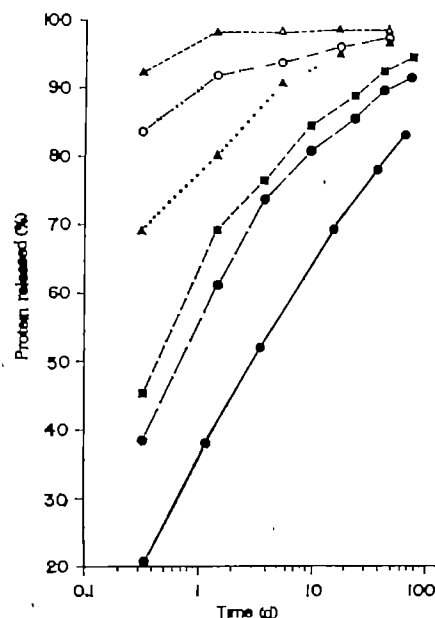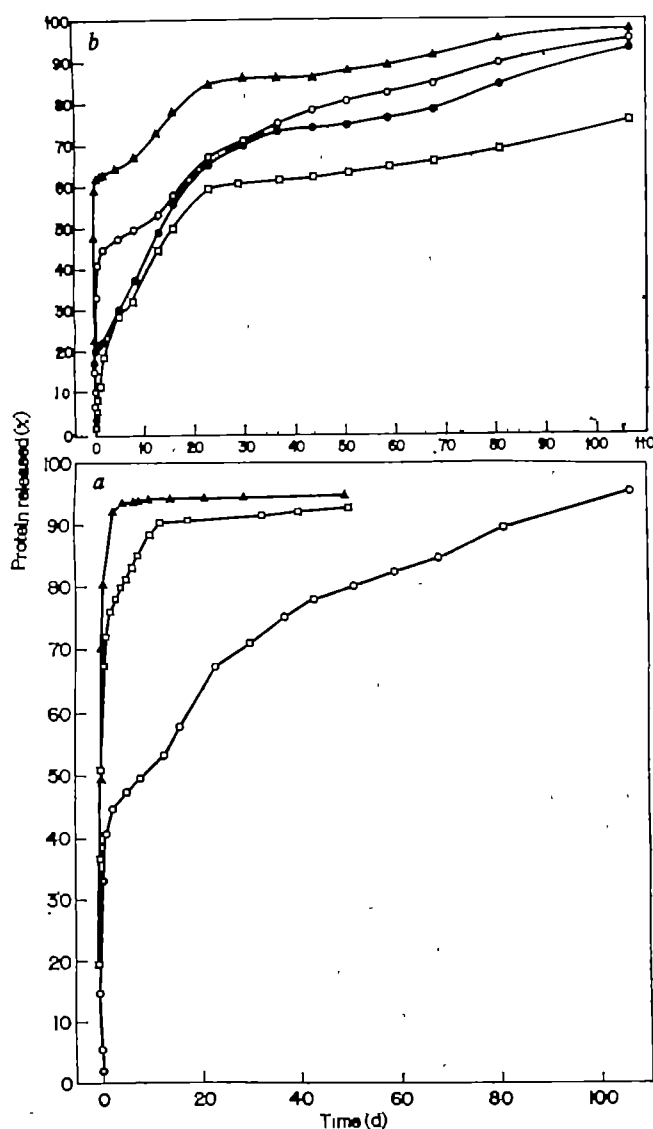
Fig. 2 a, Release of soybean trypsin inhibitor from Hydron (△), polyvinylalcohol (□) and ethylene–vinyl acetate copolymer 'sandwiches' (○). Protein concentration in all casting solutions was 50 mg ml⁻¹. Polymer concentrations in casting solutions were 12% Hydron, 10% polyvinylalcohol and 10% ethylene–vinyl acetate copolymer. These concentrations minimised initial rates of release from each of the individual polymer systems and still permitted viscosities low enough so that the solutions were easy to work with. b, Release of four proteins from ethylene–vinyl acetate copolymer 'sandwiches'. Protein and polymer concentrations in casting solutions were 50 mg ml⁻¹ and 10%, respectively. Release was studied as described in Fig. 1. △, Lysozyme; ○, soybean trypsin inhibitor; ●, alkaline phosphatase; □, catalase.

Three other proteins, lysozyme (molecular weight 14,400), alkaline phosphatase (molecular weight 88,000), and catalase (molecular weight 250,000) were similarly studied. Their release profiles from ethylene–vinyl acetate copolymer are shown in Fig. 2b. Release rates approach zero-order kinetics over periods of 20–100 d. Using standard assays[7–9], we also found that this polymer released more than 100 ng per day of heparin, and ³H-thymidine-labelled DNA (isolated from mouse SVT2 cells); and in excess of 100 μU per day of insulin for periods exceeding 40 d.

To determine whether the polymers were releasing biochemically active material in vitro, pellets containing 300 μg of lysozyme, soybean trypsin inhibitor, and alkaline phosphatase were placed in 20-ml volumes of Ringer at 37 °C. After 2 d of this incubation, each polymer was washed with 100 ml of Ringer for 5 min and placed on an agar diffusion

slide described before[10–12]. Formation of zones of activity on these slides indicated the release of at least 100 ng per day of biochemically active material. When this cycle was repeated, new zones continued to form for periods exceeding 100 d with ethylene–vinyl acetate copolymer 'sandwiches', and for 14 and 4 d with polyvinylalcohol and Hydron 'sandwiches', respectively. Furthermore, when the specific activity of alkaline phosphatase was tested before and after incorporation into each of the three polymers, more than 80% of the enzyme escaping from the polymers was found to be active.

To determine whether biologically active macromolecules could be released in vivo, ethylene–vinyl acetate copolymer 'sandwiches' containing tumour angiogenesis factor (TAF), a tumour cell extract that induces neovascularisation[13], were tested. Pellets containing TAF were implanted into 20 rabbit corneas and in every case vessels sprouted from the corneal edge and grew towards the polymer. No oedema or white cells were detected. After 40 d, several of the polymers were removed and vessels disappeared within 21 d. When the same polymers were washed in Ringer solution and transplanted to new corneas, vascular responses were again observed, although rates of vessel growth were somewhat slower than with the initial implants. Also, pellets of polyvinylalcohol or ethylene–vinyl acetate copolymer containing TAF consistently produced neovascular responses when implanted on a 10- or 11-d-old chick chorioallantoic membrane. The polymers alone induced no such response.

These studies show that sustained release of proteins and other macromolecules from polymeric vehicles can be achieved over prolonged periods. The mechanism of release of large molecules from these polymers is not entirely clear. The mechanism, at least for ethylene–vinyl acetate copolymer, does not seem to depend on simple diffusion. The same macromolecules which were released so readily from this polymer after incorporation into its matrix, do not diffuse through a film of the pure polymer. Furthermore, when Dextran blue (molecular weight 2,000,000) was incorporated into ethylene–vinyl acetate copolymer, more than 90% was released in 3 d, in contrast to proteins such as alkaline phosphatase (molecular weight 88,000) where 90% release required nearly 100 d. It can be speculated that because this polymer is permeable to water, proteins deep within the polymer matrix may be driven to the surface by osmotic pressure or capillary action.

There are several possible applications. These polymers may be useful in the development of bioassays for various substances, including growth factors, chemotactic compounds and other informational molecules. The ability of ethylene–vinylacetate copolymer in casting solution to bind to plastics may facilitate new approaches for tissue culture experiments. It remains to be seen whether these polymers will eventually be clinically useful for long term delivery of macromolecules such as insulin, heparin or enzymes.

ROBERT LANGER
JUDAH FOLKMAN

Department of Surgery,
Children's Hospital Medical Center and
Harvard Medical School,
Boston, Massachusetts 02115

1 Folkman, J., and Long, D. M., J. surg. Res., 4, 139–142 (1964).
2 Davis, B. K., Experientia, 28, 348 (1972).
3 Davis, B. K., Proc. natn. Acad. Sci. U.S.A., 71, 3120–3123 (1974).
4 Gimbrone, M. A., Jr, Cotran, R. S., Leapman, S. B., and Folkman, J., J. natn. Cancer Inst., 42, 314–427 (1974).
5 Fromer, C. H., and Klintworth, G. K., Am. J. Path., 79, 537–554 (1975)
6 Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J., J. biol Chem. 193, 265–275 (1951).

7 Bittner, T , and Muir, H , *Analyt. Biochem* , 4, 330–334 (1962).
8 Furano, A V., *Analyt Biochem* , 43, 639–640 (1972)
9 Soeldner, J S , and Slone, D , *Diabetes*, 14, 771 (1965).
10 Carlsson, L , and Karlsson, B , *Experientia*, 28, 990–991 (1972)
11 Schill, W. R., and Schumacher, G F. B , *Analyt. Biochem* , 46, 502–533 (1972)
12 Schumacher, G F B , and Schill, W R , *Analyt. Biochem* , 48, 9–26 (1972).
13 Folkman, J , *Cancer Res* , 34, 2109–2113 (1974)

Enhancement in the sweetness of sucrose

THERE IS considerable interest in the relationship between chemical structure, both configurational and conformational and the sweet response[1] particularly in the case of sucrose (1) since, of all the sweetening agents it is the most widely used Hitherto, no derivative of sucrose nor any other carbohydrate has shown sweetness substantially greater than that of sucrose, indeed to the contrary sucrose octa-acetate is very bitter, sucrose monoesters are much less sweet and *galacto*-sucrose (2), a C4 epimer of sucrose, is greatly diminished in sweetness[2]. Nevertheless, one of the objectives in our studies on the chemistry of sucrose has been to modify its structure, by chemical means in order to enhance its natural sweetness and in the process prevent metabolism by inhibiting breakdown by invertase. We have now been greatly encouraged by the surprising discovery that a *galacto*-sucrose derivative, namely 1',4,6,6'-tetrachloro-*galacto*-sucrose (3) is intensely sweet, comparable with saccharin but without an unpleasant after-taste.
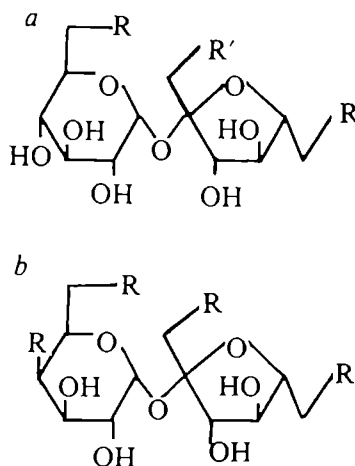
Fig. 1 *a*, Sucrose (1; R = R' = R'' = OH) and its 6,6'-dichloro (4; R = R'' = Cl, R' = OH) and 1',6'-dichloro (5; R' = R'' = Cl, R = OH) derivatives; *b*, *galacto*-sucrose (2; R = OH) and its tetrachloro derivative (3; R = Cl)

Lindley, Birch and Khan[3] suggested from studies on methyl ethers of sucrose that the equatorial 4-hydroxyl group of sucrose (1) interacts with the taste buds of the tongue by intermolecular hydrogen bonding and that loss of sweetness in *galacto*-sucrose (2) is due to the axial 4-hydroxyl group being masked by an intramolecular hydrogen bond to the ring oxygen. Whilst our tetrachloro derivative (3) is not in disagreement with the latter observation, the 4-chloro substituent could only act as a hydrogen bond acceptor on the taste bud. Likewise they suggested the importance of the 1'-hydroxyl group of sucrose which is supported by our own observations in that 6,6'-dichloro-sucrose[3] (4) is less sweet than sucrose, implying that in (3) the 1'4-dichloro substituents are enhancing its sweetness. On the other hand 1',6,6'-trichloro[4] and 4,6,6'-trichloro[5] derivatives are more than 10 times sweeter than sucrose but not as intensely sweet as the aforementioned tetrachloro derivative (3). It is highly significant that 1'6'-dichlorosucrose (5) is intensely sweet, like the tetrachloro derivative (3), which

pinpoints the importance of the chloro substituent at the 1'-position on the molecule of sucrose.

If the Shallenberger theory of sugar sweetness[6] is correct, where two electronegative atoms A and B separated by 2.5–4 0 Å initiate the sweet taste by intermolecular hydrogen bonding between A–H and B and similar groups on the receptor site, then it would seem that the 1'-chloro substituent is the proton acceptor B and the 2-hydroxyl on the glucosyl moiety is A–H, the proton donor. The intense sweetness of the chloro compounds 3 and 5 could also be due in part to their greater lipophilic character than sucrose. The subtle interplay of intramolecular bonding in the sucrose molecule has an important role and the introduction of chloro substituents seems to have reinforced the effect. The sucrose derivatives in which a 6-hydroxyl of the fructo-furanosyl unit of sucrose has been replaced by a chloro substituent are resistant to the action of invertase, a β-D-fructofuranosidase (J. Chapman, unpublished results). Hence in the development of alternative sweeteners from sucrose, as for example for application in dietetic food, combining high sweetness with low calorie content, the chloro derivatives are worthy of further evaluation.

L. HOUGH
S. P. PHADNIS

Department of Chemistry,
Queen Elizabeth College,
(University of London),
London W8 7AH, UK

1 Birch, G G , Green, L F., and Coulson, C B , *Sweetness and Sweeteners* (Applied Science Publishers, London, 1971)
2 Lindley, M G , Birch, G. G., and Khan, R., *J Sci Fd Agric.*, 27, 140–144 (1976).
3 Hough, L , and Mufti, K S , *Carbohyd Res* , 25, 497–503 (1972)
4 Khan, R., *Carbohyd Res* , 25, 504–510 (1972)
5 Hough, L , Phadnis, S P , and Tarelli, E , *Carbohyd Res* , 44, 37–44 (1975)
6 Shallenberger, R S., and Acree, T E , *Nature*, 216, 480–482 (1967)

Erratum

In the article "Conformation of met[5]-enkephalin determined by high field PMR spectroscopy" by B. P Roques *et al.* (*Nature*, 262, 778; 1976) the legends to Figs 2 and 3 are transposed. The Figs are in the correct positions The alignment of the stereo picture is incorrect and should be as below.
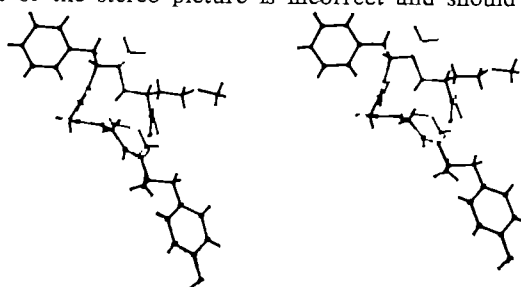
reviews

Tackling computer myths
Yorick Wilks

The Thinking Computer: Mind Inside Matter. By Bertram Raphael. Pp. xiii + 322. (Freeman: San Francisco and Reading, May 1976.) Cloth £9; paper £4.40.

ARTIFICIAL INTELLIGENCE (AI) is beset by all the difficulties of any immature science. Since so little is agreed among its workers in the way of basic theory and truths, any presentation of it is bound to be inadequate, and that applies especially to any attempt to produce a non-contentious introductory text like the one under review. It is almost inevitable that such a work will be parochial in terms of the research laboratory it comes from, and the areas within the field on which it concentrates. To say this is not to blame the author: in the present state of things in AI it could not be otherwise.

Thus the present book is heavily oriented to the research carried out at the Stanford Research Institute (SRI), its author's home, and is strongest in the areas where SRI is strong: robots, formal approaches to problem solving and vision; and weakest where they are weakest: natural language, high level programming languages, and the sort of speculation about 'knowledge structures' that has been emerging from MIT in the past two years. The book seems to have been written essentially by about 1972, and is, in that sense, out of date. But, once again, it is a criticism of the subject and not the author that being up to date is considered so important in AI. In a mature science, the basics could be absorbed perfectly well from a four-year-old text-book. In so far as abstract problem-solving techniques, resolution theory in theorem proving, and techniques of the same vintage in vision, are fundamental building blocks of the subject (and to some extent it is undeniable that they are), then they are well and clearly presented here.

The book aims to fill the gap between a wholly popular account of AI research and the style of the technical literature; and is intended to lead anyone interested and armed with 'High School' maths to that literature. It does it rather well, and could be a useful book for first-year University courses in computer science or psychology.

Its concessions to laymen are among its best bits. Raphael begins by tackling what he calls "myths about computers": the Arithmetic Myth and the Stupid Computer Myth, namely that computers can do only what they are programmed to do, which is true but highly misleading. While he was at it he should probably have tackled the contrary myths: that computers are socially all-powerful and as clever as we are. The latter is just as likely to be believed by the man in the street at the moment as the "myths" Raphael does tackle. The book's subtitle, too, *Mind Inside Matter*, is peculiarly unfortunate for an author concerned to debunk myths.

Another bonus is the well illustrated chapter on the history of robots, which contains the marvellous revelation (p 253) that "several key figures in contemporary computer history"—Von Neumann, Wiener and Minsky—are all descendants of the Maharal· the fifteenth-century rabbi of Prague who made, and gave life to, a 'golem', an artificial man made from the clay of a river bank. □

Yorick Wilks has worked on Artificial Intelligence and natural language understanding at Cambridge, Stanford and in Switzerland. He is at present a Senior Visiting Fellow at the University of Edinburgh, UK.

Matrix isolation

Matrix Isolation: A Technique for the Study of Reactive Inorganic Species. By Stephen Cradock and A. J. Hinchcliffe. Pp. 144. (Cambridge University: Cambridge, London and New York, October 1975.) £6.20.

Now acknowledged as one of the primary methods of investigating molecules which are short-lived in conventional conditions, matrix isolation has not lacked reviewers, but books about it have been few and relatively specialised. The present authors are to be commended therefore on producing a compact and eminently readable account of matrix isolation which can easily be digested by undergraduate and graduate students unfamiliar with the technique. The formation and nature of a matrix, interactions within that matrix, and the technical aspects of matrix-isolation experiments are treated at a simple, descriptive level, specific examples are discussed in the closing chapters. There is a short subject index and a cameo-sized bibliography that is bound to disappoint anyone with more than passing interest in the subject.

Assuming that the reader is familiar with much basic physics, chemistry and spectroscopy, the authors have aimed to show how matrix-isolation experiments may be accomplished Assuredly, theirs is a reasonable objective, but the result seems to me overburdened with empiricism, and addicted to assertions unsupported either by argument or by apt example. Without a broader appreciation of the purpose and possibilities of the technique, the reader might be excused for wondering whether the technical mountain has not given birth to a chemical mouse. This doubt is not altogether dispelled either by the range or by the treatment of the examples selected. Certain it is that the authors take a healthily critical view of the results of matrix-isolation experiments: whether those results signify much outside their immediate context has to be taken on trust. If the figures would have repaid a greater expenditure of imagination and trouble, the way in which the story is told recalls Mark Twain's character: "there was things which he stretched, but mainly he told the truth" Mainly this is a sound and worthwhile addition to the chemical literature: it might have been more than that.

A. J. Downs

Dr Downs is a Lecturer in the Department of Inorganic Chemistry at the University of Oxford, UK.

• Catecholamine
· transport systems

The Mechanisms of Neuronal and Extraneuronal Transport of Catecholamines. Edited by David M. Paton. Pp. xi+370. (Raven: New York, February, 1976.) $27.50.

THE discovery by Axelrod and his colleagues that radioactively labelled catecholamines are transported into sympathetically innervated tissues has led to a much better understanding of the physiology and pharmacology of adrenergic neurotransmission, and the early work was summarised in 1967 in what is now becoming a classic book *The Uptake and Storage of Noradrenaline in Sympathetic Nerves* by L. L. Iversen. It is particularly welcome, therefore, that we are brought up to date by the book under review. The editor, in an historical introduction, also reminds the reader that possibly the earliest suggestion that circulating adrenaline might be transported into nerve endings was made by J. H. Burn

in 1932 and first shown experimentally in 1943 by Raab.

The book contains 17 chapters contributed by well-known workers in the field and covers the major catecholamine transport systems known, those across the membranes of sympathetic neurones, of synaptic vesicles and of non-neuronal cells in a variety of tissues. It is of fairly specialised interest, but most workers who study catecholamines and other neurotransmitters will want a copy in their library. Unfortunately, there are some inconsistencies in the book between different authors. Several define precisely what is meant by the term 'uptake' and their definition is welcome; but other authors in the book use the term loosely and are less critical. Finally, the title was not a good choice, for only one chapter covers what is generally understood by the word 'mechanism': the way in which the catecholamine molecule passes across a membrane.

A. D. Smith

Dr Smith is a Lecturer in the Department of Pharmacology at the University of Oxford, UK.

Spectroscopy
updated

Spectroscopy. Vol. 1 : Pp. 304. Vol. 2 : Pp. x+362. Vol. 3 : Pp. x+324. (Science Paperbacks.) Edited by B. P. Straughan and S. Walker (Chapman and Hall: London; Halsted · New York, June 1976.) Paper £5.50 each volume; hardback £9 each volume.

THIS set of three volumes consists of a major reform of *Spectroscopy* by S. Walker and H. Straw issued in two volumes in 1961–62. It covers most of non-nuclear spectroscopy with an approach favouring the molecular aspects rather than deep fundamental theory. Each of the nineteen chapters has its individual authorship, although in some places recognisable parts of the earlier work re-appear. There is a modest amount of cross reference, a degree of common style and no serious duplication apart from the deliberate repetition of an appendix containing some more mathematical features in both volumes 2 and 3. The editors have been effective in imposing SI where units are required but they have not had an eagle eye on the associated algebraic forms, many of which lack a factor $(1/4\pi\varepsilon_0)$ to be consistent with SI They also eschew h in favour of the cumbersome $h/2\pi$.

Volume 1 contains· atomic spectroscopy and various forms of spin resonance including quadrupole nuclei and the Mössbauer effect. In this volume,

nuclear magnetic resonance is only allocated 64 pages which seems light for balance and solids are largely omitted. Volume 2 contains the rotation and vibration features and a chapter on group theory which is at a deeper level than is common in such student texts; the chapter on force constants is more appropriate to this scientific level. The chapter on the far infrared provides interesting experimental detail; instrumentation elsewhere is inevitably briefly covered. Volume 3 covers essentially electronic molecular spectroscopy and includes a special chapter on photoelectron spectroscopy and another on astrophysics and astrochemistry.

This replacement is timely and will be widely welcomed by those who found the earlier edition valuable. The books are somewhat lengthy in parts for chemistry or physics undergraduates, but the volumes should be available in the libraries they use. At M Sc level— or possibly undergraduate chemical physics—it would be very suitable; and also as an introduction to research. It can be strongly recommended for these purposes, although it retains a slightly old fashioned approach to quantum theory which restrains the enthusiasm of this reviewer   **D. H. Whiffen**

Professor Whiffen is Head of the Department of Physical Chemistry at the University of Newcastle upon Tyne, UK.

Hydrolysis
equilibria of
metal cations

The Hydrolysis of Cations. By Charles F. Baes, Jr, and Robert E. Mesmer. Pp. xxi+489. (Wiley-Interscience: New York and London, May 1976.) $49.40; £27.66.

THE first three chapters of this book will be useful to anyone who wishes to acquire a good background knowledge before beginning a study of the complicated hydrolysis equilibria of metal cations. Methods of measurement (chiefly potentiometric) and the interpretation and treatment of data are covered systematically and in a clear style. The authors conclude this section of the book (pp69–70) with a slightly philosophical attempt to answer the question: How certain can we be that the scheme of equilibria that "best" fits the data is indeed the correct one? The meaning to be attached to the word "correct" in this question would seem to be open to a lengthy discussion among philosophers of science.

The next fourteen chapters are devoted to accounts of specific cations, covering almost all the appropriate elements in the periodic table. This part of the book fulfils the authors' object of assembling in convenient form a great deal of information, most of which was published during the 25 years up to 1974. It is not claimed that the information on any particular topic is complete, but each section (together with the references—there are about 900 altogether) provides a good starting point for further study. In the last chapter a brave attempt is made at the difficult task of surveying and rationalising the variety of behaviour described earlier in detail.

There is no conventional subject index but there is a single-page alphabetical index of hydrolysing species, a chapter index in the form of a periodic table and a ten-page list of contents. The book should therefore be quite convenient to use, but the casual reader should not overlook the summary of symbols and definitions. These are not always conventional, and they include the rather unexpected definition. $pH = -\log[H^+]$. The standard of production is high—as is the price. It is a book for libraries rather than for the individual buyer.   **J. S. Coe**

Dr Coe is a Senior Lecturer in the Department of Chemistry at King's College, University of London, UK.

Electron donor–acceptor complexes

Catalysis by Electron Donor–Acceptor Complexes: Their General Behaviour and Biological Rôles. By Kenzi Tamaru and Masaru Ichikawa Pp viii+208 (Kodansha: Tokyo; Wiley · New York and London, March 1976.) $23.40; £11.70

IN view of the large number of recent books and reviews dealing with various aspects of electron donor–acceptor (EDA, or charge-transfer) complexes, one must ask whether this book contributes anything new Unfortunately, aside from one chapter, the answer is no.

The central theme of the book is the chemical properties of EDA complexes, including their reactions, catalysis by EDA complexes, and the biological importance of EDA complexes. Chapters 1 (Introduction, 8 pages) and 2 (Formation of EDA Complexes, 44 pages) present a brief general introduction to the subject of EDA complexes Most of this material is covered more clearly and accurately in earlier books and reviews.

Chapter 3 (Homogeneous Catalysis by EDA Complexes, 40 pages) discusses reactions of EDA complexes and catalysis by complexing in solution. This chapter contains numerous inaccurate and confusing statements and is neither complete nor up to date. Again, this subject is covered much more adequately in any one of several recent books and reviews.

Chapter 4 (Heterogeneous Catalysis by EDA Complexes, 66 pages) deals with the subjects of adsorption through EDA complexing and by EDA complexes, and catalysis by solid EDA complexes. This topic, being the main interest of the authors, is covered in much more detail with numerous references to recent work. I found the chapter interesting and relatively free from errors. No comparable review of these subjects is available elsewhere

Chapter 5 (Role of EDA Complexes in Biochemical Reactions, 40 pages) is again neither complete nor up to date and the subject is covered more adequately elsewhere.

A very substantial fraction of Chapter 5 as well as smaller parts of Chapters 2 and 3 are copied almost verbatim from earlier books and reviews. Perhaps this is why, with the exception of Chapter 4, there are few references later than 1968.

In summary, the only very useful part of this book is the chapter dealing with heterogeneous catalysis Only persons actively interested in this area

are likely to consider paying $23.40 for a 66-page review on the subject

Allan K. Colter

Allan K. Colter is Professor of Chemistry at the University of Guelph, Guelph, Ontário, Canada

Vibrational spectroscopy

Vibrational States. By S. Califano. Pp. xii+335. (Wiley–Interscience· London and New York, April 1976.) £16.75; $34.50.

THIS is a textbook on the theory of vibrational spectroscopy for graduate students. It evolved from a lecture course given by the author in the University of Florence, and is restricted to the theoretical treatment of vibrations of isolated molecules.

The book consists of nine chapters, plus two appendices containing group-theoretical character and correlation tables Chapter 1 is a short introduction to infrared and Raman spectroscopy. Chapter 2 presents the classical theory of small molecular vibrations and includes a very clear account of the separation of translational, rotational and vibrational motion (the Eckart or Sayvetz conditions). The quantum theory is given in Chapter 3, it deals with the one-, two- and three-dimensional harmonic oscillator and with the quantum radiation field. Chapter 4 is concerned with the transformation to internal coordinates and the construction of the F and G matrices for the

force constants and the kinetic energy in terms of the internal coordinates Chapters 5 and 6 are devoted to the principles of group theory and its applications to molecular vibrations, and Chapter 7 to selection rules in infrared and Raman spectra. Various potential functions, including the general valence force field and the Urey–Bradley force field, are discussed in Chapter 8 which includes a brief description of Coriolis coupling, centrifugal distortion, and the *ab initio* computation of force constants. The final chapter is concerned with anharmonicity and Fermi resonance.

The author has given us a clear and detailed account of the main topics in the theory of molecular vibrations. There are omissions, including vibronic coupling, internal rotation, circular dichroism, and Raman spectra resulting from the anti-symmetric part of the polarisability tensor; and a few obscurities and minor errors, such as the comparison at the foot of p20 of quantities having different dimensions. SI units are not used, nor are Mulliken's recommendations (*J chem. Phys.*, **23**, 1997, 1955) for the choice of molecule-fixed axes; the coordinate frames on p19 are left-handed and some of the diagrams (such as that showing a plane of symmetry on p103) are not clear. But overall this is a good book that will be a useful addition to the well-known texts by Herzberg and by Wilson, Decius and Cross.   **A. D. Buckingham**

A D. Buckingham has been Professor of Chemistry at the University of Cambridge since 1969.

Environmental dictionary

Dictionary of Environmental Terms. By Alan Gilpin. Pp. 191. (Routledge and Kegan Paul. London and Henley, September 1976.) £3.50.

THIS book is attractively produced, easy to use, and has some very useful features: the list of abbreviations at the beginning (although why does the MW=megawatt come between p.p.m. and its expansion?) and the Appendix carrying the Statement of the Stockholm Conference on the Human Environment (1972). The important criteria in reviewing such a book are: "Does it contain puzzling words that the student/layman is likely to encounter?" and "Are the definitions sound?". The range of words is admirably broad, but there are some omissions, especially of

biological terms. No 'biotic index'. 'sere' or 'ecotype', for instance, and only 'meteorological stability'. On the pollution side 'fanning' of plumes appears, but no 'slug' or 'slick'; and although 'effluent charge' is included 'consent' or 'consent conditions' is not

'Qualitative and quantitative analysis' probably need not have been included, and the long description of Boulding's green stamp plan for creating a market in procreation, although intriguing, is of only marginal importance.

Some of the explanations are too discursive, more appropriate to a gazetteer than a dictionary. In spite of these criticisms, however, this is a helpful book reasonably priced.

Pauline K. Marstrand

Pauline Marstrand is a Senior Fellow in the Science Policy Research Unit, University of Sussex, UK.

obituary

With the death of **Professor J. C. Slater** at the age of 75, America, and indeed the world, has lost one of its most influential and active practitioners of theoretical physics. For over fifty years Slater produced a continuous and substantial flow of research work, most of it concerned with those aspects of matter which depend primarily on the electronic properties of atoms; treating first the properties of individual atoms as revealed by the multiplet structure of their spectra, and later the properties of atoms in molecular combinations and in the solid state

With the advent of wave mechanics he quickly realised that all such properties are determined by the many-electron wave function of the system, and therefore directed his attention to the invention of suitable approximations and mathematical methods for its calculation. In his search for the wave functions, their eigenvalues, and the density of states they imply, he led a numerous band of gifted students many of whom (now somewhat advanced in years) have followed the trail which he blazed and, with more powerful computers and more refined techniques, have achieved even better results Surely no greater tribute can be paid to a scientist than the successful continuation of his work by succeeding generations.

John Clarke Slater was born on December 22, 1900 into an academic family; his father was professor of English at the University of Rochester. John, after taking his bachelor's degree at this university, went on to Harvard as a graduate student to work under the direction of P W Bridgman In 1922 he travelled to Europe and during a stay of a few months in Copenhagen he became part author, with Bohr and Kramers, of a well-known paper on the quantum theory of radiation Although

the conclusions of this paper proved to be incorrect, it was influential for a time, and for a young man of 23 the association with such famous names was a happy augury indeed.

In 1929 Slater published one of his most important papers which dealt with the theory of multiplets in complex spectra. Folowing the work of F. Hund and applying the newly developed wave mechanics he was able to simplify and extend the calculations of the multiplet structure of atoms with many electrons The basic concept in this work was the expression of the total wave function as a determinant of the single electron wave functions including their spin states Such determinants have ever since been known as Slater determinants and have found applications in many fields, for example, chemistry and metal physics

Following the publication of Heisenberg's theory of ferromagnetism and Bloch's famous paper on the quantum mechanics of electrons in crystal lattices, Slater wrote a paper in 1930 entitled *Cohesion in Monovalent Metals* The significance of this paper in Slater's scientific life is that it marks his introduction into what is now known as solid state physics and which was to remain his principal interest for over forty years. As with many other physicists, however, the years of the second world war caused an interruption in Slater's academic research. During these years he was active in organising the Radiation Laboratory at M.I.T and was involved with research concerning the principles of magnetron design For this work he was awarded a Presidential Certificate of Merit

After the war Slater returned to his interests in solid state physics and began a series of calculations aimed at improving the accuracy of wave functions and the energy levels of electrons

in metals. He developed the method introduced by Wigner and Seitz so that it could be applied to states of higher energy and used the resulting eigenvalue spectrum to construct curves of the density of states. Later he introduced the augmented plane wave method which was widely used by his own students and other physicists. This work, carried out mainly in the decade 1950–60, is not to be judged by the accuracy, or inaccuracy, of particular results but by the thrust which it gave to the advance in this sector of the subject.

Slater was a most prolific writer. Besides about one hundred original papers he wrote eleven books, a few in collaboration with N. H. Frank, but mostly alone, and in addition produced many survey articles, including one giving a detailed history of the development of solid-state and molecular theory. A notable feature of Slater's pedagogic writings is the extraordinary complete bibliography which is invariably attached to each book or article. Evidently very little escaped his attention. It seems clear that he wrote with great facility and speed and that he maintained a confidence in his knowledge and understanding of physics throughout his life

Other aspects of Slater's life and character, for example, his administrative ability which enabled him to take over the headship of the physics department of M.I.T. at the early age of 29, and his gifts as a teacher, can be dealt with adequately only by a close colleague, and in an introductory passage to the volume of articles published to commemorate his 65th birthday, Professor Philip Morse pays a warm tribute to Slater with regard to exactly these qualities. Clearly, Slater was a man of quite exceptional stature in many respects.   **H. Jones**

announcements

Meetings

March 28–30, **Nuclear Magnetic Resonance in Biology**, Oxford (Dr R A. Dwek, Department of Biochemistry, University of Oxford, South Parks Road, Oxford OX1 3QU, UK).

April 4–6, **Microcalorimetry in Biology**, London (3rd International Symposium

on Microcalorimetry, LKB Instruments Ltd., LKB House, 232 Addington Road, Selsdon, South Croydon, Surrey CR2 8YD, UK).

April 25–27, **Bioengineering**, Fort Collins (Deadline for abstracts January 1) (Dr C W. Miller, Department of Physiology and Biophysics, Collaborative Radiological Health Laboratory, Colorado State University,

Fort Collins, Colorado 80521).

April 27–29, **Food, Fertilizer and Agricultural Residues,** Syracuse, New York (Waste Management Conference, Cornell University, 207 Riley-Robb Hall, Ithaca, New York 14853).

May 21–25, **Carboniferous Stratigraphy and Geology**, Urbana–Champaign (Dr Mackenzie Gordon, U S. Geological Survey, Washington D.C.).

nature

Author Index for Volume 263

September–October 1976
Numbers 5572–5580
Pages 1–804

Recent scientific and technical books

Mathematics

BERLINSKI, David On Systems Analysis An Essay Concerning the Limitations of Some Mathematical Methods in the Social, Political and Biological Sciences Pp x + 186 ISBN-0-262-120-X (Cambridge Mass and London The MIT Press, 1976) $15

COLTON D L Solution of Boundary Value Problems by the Method of Integral Operators (Research Notes in Mathematics—6) Pp 148 ISBN-0-273-00307-0 (London San Francisco and Melbourne Pitman Publishing, 1976) £5 60

GILBERT, R P and WEINACHT, R J Function Theoretic Methods in Differential Equations (Research Notes in Mathematics—8) Pp 309 ISBN-0-273-00306-2 (London San Francisco and Melbourne Pitman Publishing, 1976) £9 50.

HALL G and WATT, J M (edited by) Modern Numerical Methods for Ordinary Differential Equations Pp 336 ISBN-0-19-853348-9 (Oxford Clarendon Press London Oxford University Press, 1976) £9 75 net

HOWIE, J M An Introduction to Semigroup Theory (LMS Monographs Vol 7) Pp x + 272 ISBN-0-12-75-46333 (London and New York Academic Press, a Subsidiary of Harcourt Brace Jovanovich Publishers, 1976) £9 80

LAPIDUS L, and SCHIESSER W E Numerical Methods for Differential Systems Recent Developments in Algorithms Software and Applications Pp x + 291 ISBN-0-12-4366406 (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich Publishers 1976) £3 50, £8 25

MONOD-HERZEN, Gabriel 1 Analyse Dimensionnelle et l'Épistémologie (Recherches Interdisciplinaires) Pp 129 ISBN-2-224-00296-3 (Paris Maloine S A et Doin S A, 1976) 68 F Fncs

POSTON, T and STEWART I N Taylor Expansions and Catastrophes (Research Notes in Mathematics—7) Pp 168 ISBN-0-273-00964-8 (London San Francisco and Melbourne Pitman Publishing 1976) £5 90

Astronomy

BOHME, S ESSER U FRICKE W GUNTZEL-LINGNER, U, HEINRICH, I, HENN, I, KRAHN D, SCHMADEL, L, SCHOLL H, and ZECH, G (edited by) Astronomy and Astrophysics Abstracts Vol 15/16 Author and Subject Indexes to Volumes 1-10 Literature 1969-1973 (A Publication of the Astronomisches Rechen-Institut Heidelberg Member of the Abstracting Board of the International Council of Scientific Unions) Pp v + 655 ISBN-3-540-07605-X (Berlin and New York Springer-Verlag, 1976 Published for Astronomisches Rechen-Institut) DM 86, $35 30

MITTON, S Exploring the Galaxies Pp 206 ISBN-0-571-10925-X (London Faber & Faber, 1976) £4 75

MUIRDEN, J Astronomy with Binoculars Pp 173 ISBN-0-571-04903-6 (London Faber & Faber, 1976) Cloth £3 95 Paperback £2 75

SHEN, B S P, and MERKER M (edited by) Spallation Nuclear Reactions and Their Applications (Astrophysics and Space Science Library, Vol 59) Pp vii + 235 ISBN-90-277-0746-X (Dordrecht and Boston, USA D Reidel Publishing Company, 1976) Dfl 75, $27 50

Physics

ARNOWITT Richard and NATH Pran (edited by) Gauge Theories and Modern Field Theory (Proceedings of a Conference held at Northeastern University, Boston, September 26 and 27, 1975) Pp 428 ISBN-0-262-01046-1 (Cambridge Mass and London The MIT Press 1976) £10 20

DEAN, Nathan W Introduction to the Strong Interactions Pp xiv + 377 ISBN-0-677-02750-8 (New York, London and Paris Gordon and Breach Science Publishers, 1976) £19 70

DULEY W W CO, Lasers Effects and Applications (Quantum Electronics—Principles and Applications a Series of Monographs) Pp xin + 427 ISBN-0-12-223350-6 (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich Publishers, 1976) $31 £18 40

FONER, Simon (edited by) Magnetism Selected Topics Pp xx + 748 ISBN-0-677-15390-2 (New York, London and Paris Gordon and Breach Science Publishers, 1976) £44

FREIFELDER, D Physical Biochemistry Applications to Biochemistry and Molecular Biology Pp x + 570 ISBN-0-7167-0560-5 (San Francisco W H Freeman and Company, 1976) Hardcover £12 40 Softcover £8 00

GRUNDY, P J and JONES, G A Electron Microscopy in the Study of Materials Pp 174 ISBN-0-7131-2521-7 (London Edward Arnold, 1976) Boards £8 00, Paperback £3 75

HARVEY A F (compiled by) Bibliography of Microwave Optical Technology (Solid State Physics Literature Guides, Vol 8) Pp xxxi + 717 ISBN-0-306-68328-8 (New York, Washington and London IFI/Plenum, 1976) $60 69

MAURIN Krzysztof, and RACZKA, Ryszard (edited by) Mathematical Physics and Physical mathematics (Proceedings of the International Symposium held in Warsaw, 25-30 March 1974) Pp xvii + 504 ISBN-90-277-0537-2 (Dordrecht and Boston USA D Reidel Publishing Company, Warsaw PWN-Polish Scientific Publishers, 1976) Dfl 105, $39

PAUL, E, ROLLNIK H, and STICHEL, P Elementary Particle Physics (Springer Tracts in Modern Physics, Vol 79) Pp vii + 145 ISBN-3-540-07771-2 (Berlin and New York Springer-Verlag 1976) DM 64, $26 30

PERLMUTTER, Arnold (edited by) New Pathways in High-Energy Physics I Magnetic Charge and Other Fundamental Approaches (Studies in the Natural Sciences, Vol 10) Pp x + 401 ISBN-0-306-36910-9 (v 1) (New York and London Plenum Press, 1976) $47 40

PERLMUTTER, Arnold (edited by) New Pathways in High-Energy Physics II—New Particles, Theories and Experiments (Studies in the Natural Sciences) Pp x + 411 ISBN-0-306-36911-7 (v 2) (New York and London Plenum Press, 1976) $47 40

PINTAR, M M (edited by) Introductory Essays (NMR Basic Principles and Progress/Grundlagen und Fortschritte, Vol 13) Pp xi + 154 ISBN-3-540-07754-5 (Berlin and New York Springer-Verlag, 1976) DM 72, $29 60

ROSENFELD, A (edited by) Digital Picture Analysis (Topics in Applied Physics Vol 11) Pp xin + 351 ISBN-3-540-07579-8 (Berlin and New York Springer-Verlag, 1976) DM 72, $29 60

SHEFTAL, N N (edited by) Growth of Crystals, Vol 10 Pp x + 290 ISBN-0-306-18110-X (Translated by J E S Bradley (New York and London Consultants Bureau, 1976) $45

STEINFELD, Jeffrey I Electronic Transition Lasers (Proceedings of the Second Summer Colloquium held at the Marine Biology Laboratory Conference Center, Woods Hole, Massachusetts, 17-19 September, 1975) Pp xi + 311 ISBN-0-262-19146-6 (Cambridge Mass and London The MIT Press, 1976) $14 95

THE CHEMICAL SOCIETY A Specialist Periodical Report Surface and Defect Properties of Solids Vol 5 Pp ix + 233 ISBN-0-85186-290-X, (London The Chemical Society, 1976) £16.00,

WEISS, Richard A Relativistic Thermodynamics, Vol 1 Pp x + 94 ISBN-0-682-48532-2 (Jericho, New York Exposition Press, 1976) $10

WICHTERLE, I, LINEK, J and HALA, E Vapor-Liquid Equilibrium Data Bibliography 1968 ISBN-0-444-41161-5 (Amsterdam London and New York Elsevier Scientific Publishing Company, 1976) Dfl 100 00, $38 50

WILLIAMS, Dudley (edited by) Spectroscopy, Part B (Methods of Experimental Physics, Vol 13) Pp xv + 364 ISBN-0-12-475954-8 (Pt B) (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich, 1976) $37 50 £22 85

WOLD, H W Mathematical Methods for Physics (Lecture Notes and Supplements in Physics,) Pp xvi + 628 ISBN-0-8053-9856-2 (Reading, Mass and London W A Benjamin, Inc, (Cloth $26 50 Paper $16 50

Chemistry

FRAME Jr, Edward G and GRASSELLI, Jeanette G (edited by) Infrared and Raman Spectroscopy, Part A Pp ix + 345 ISBN-0-8247-6392-0 (New York and Basel, Marcel Dekker, Inc, 1976) np

DOBBS, L 1 T The Foundations of Newton s Alchemy or "The Hunting of the Greene Lyon" Pp xv + 300 ISBN-0-521-20786-X (Cambridge, London New York and Melbourne Cambridge University Press, 1976) £10 50

EYRING Henry and HENDERSON, Douglas (edited by) Theoretical Chemistry Advances and Perspectives Vol 2 Pp xi + 298 ISBN-0-12-681902-5 (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich Publishers, 1976) $30, £18 30

FLEMING, Ian Frontier Orbitals and Organic Chemical Reactions Pp vii + 249 ISBN-0-471-01820-1 (London and New York Wiley-Interscience, John Wiley and Sons, 1976) £3 95, $8 50

GERRARD W Solubility of Gases and Liquids a Graphic Approach—Data-Causes-Prediction Pp x + 275 ISBN-0-306-30866-5 (New York and London Plenum Press, 1976) $47 40

HANNAY, N B (edited by) Treatise on Solid State Chemistry Vol 3 Crystalline and Noncrystalline Solids Pp xvi + 774 ISBN-0-306-35053-X (New York and London Plenum Press, 1976) np

HANNAY, N B (edited by) Treatise on Solid State Chemistry Vol 4 Reactivity of Solids Pp xvii + 721 ISBN-0-306-35054-8 (New York and London Plenum Press, 1976) $54

MA T S, and HORAK, V Microscale Manipulations in Chemistry (Chemical Analysis a Series of Monographs on Analytical Chemistry and Its Applications) Pp xin + 488 ISBN-0-471-55799-4 (New York and London Wiley-Interscience, John Wiley and Sons, 1976) $33, £19 40

MEARES, Patrick (edited by) Membrane Separation Processes Pp 600 ISBN-0-444-41353-9 (Amsterdam, Oxford and New York Elsevier Scientific Publishing Company, 1976) Dfl 250, $6 25

RABO Jule A (edited by) Zeolite Chemistry and Catalysis (ACS Monograph, 171) Pp viii + 796 ISBN-0-8412-0276-1 (Washington, DC American Chemical Society, 1976) $65

RADLEY, J A Examination and Analysis of Starch and Starch Products Pp vii + 220 ISBN-0-85334-692-5 (London Applied Science Publishers, Ltd, 1976) £15

SMITH A L (edited by) Theory and Practice of Emulsion Technology (Proceedings of a Symposium organized by the Society of Chemical Industry and held at Brunel University, September 16-18, 1974) Pp vi + 352 ISBN-0-12-651250-7 (London and New York Academic Press, a Subsidiary of Harcourt Brace Jovanovich, Publishers, 1976) £14 50

WARE, William R (edited by) Creation and Detection of the Excited State, Vol 4 Pp x + 320 ISBN-0-8247-6451-X (New York and Basel Marcel Dekker, Inc, 1976) SFr 130

ZIMMER, Hans, and NIEDENZU, Kurt (edited by) Annual Reports in Inorganic and General Syntheses—1975 Pp ix + 344 (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich Publishers, 1976) $18, £11

Technology

BECK, A H W Statistical Mechanics Fluctuations, and Noise (Contemporary Electrical Engineering Series) Pp xin + 231 ISBN-0-7131-3363-5 (London Edward Arnold (Publishers), Ltd, 1976) Boards £12 50, Paper £7 50

CAPPER P Leonard, and CASSIE, W Faber The Mechanics of Engineering Soils Sixth edition Pp xin + 375 ISBN-419-10990-0 (London E and F N Spon, Ltd, New York Halsted Press, a Division of John Wiley and Sons, Inc, 1976) £4 85

CLARKE Robin Technological Self-Sufficiency Pp 302 ISBN-0-571-11057-6 (London Faber and Faber, Ltd, 1976) Cloth £5 75 Paperback £2 95

DENTON Jesse C, and AFGAN, Naim H Future Energy Production Systems Heat and Mass Transfer Processes Vol 1 Pp xi + 1-378 ISBN-0-12-210001-8 (v 1) Vol 2 Pp xi + 379-866 ISBN-0-12-210002-6 (v 2) (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich Publishers, Washington and London Hemisphere Publishing Corporation, 1976) $38, £23 20 each volume

FAST J D Gases in Metals (Philips Technical Library) Pp vii + 260 ISBN-0-333-21214-2 (London and Basingstoke The Macmillan Press, Ltd, 1976) £5 95

GRAY, W A, KILHAM J K and MULLER R Heat Transfer from Flames (Topics in Fuel Science) Pp 152 ISBN-0-236-40044-4 (London and New Hampshire, USA Elek Science 1976) £3 75 net

HENLEY, Ernest J, and LEWINS Jeffery (edited by) Advances in Nuclear Science and Technology Vol 9 Pp xv + 359 ISBN-0-12-029309-9 (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich, Publishers, 1976) $36 50, £21 15

IRVINE Jr Thomas F, and HARTNETT, James P (edited by) Advances in Heat Transfer, Vol 12 Pp xiv + 296 ISBN-0-12-020012-0 (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich Publishers, 1976) $31, £18

KARAM, R A, and MORGAN, K Z (edited by) Energy and the Environment Cost-Benefit Analysis (Proceedings of a Conference held June 23-27 1975, Sponsored by the School of Nuclear Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA) Pp vii + 695 ISBN-0-08-020644-1 (Oxford and New York Pergamon Press, 1976) $50, £25

KMETZ, A R and von WILLISEN, F K (edited by) Nonemissive Electrooptic Displays Pp xin + 360 ISBN-0-30630957-2 (New York and London Plenum Press, 1976) np

MARCHELLO, J M Control of Air Pollution Sources (Chemical Processing and Engineering An International Series of Monographs and Textbooks) Pp viii + 630 ISBN-0-8247-6182-00 (New York and Basel Marcel Dekker, Inc, 1976) $49 50

MILORA, Stanley L and TESTER, Jefferson W Geothermal Energy as a Source of Electric Power Thermodynamic and Economic Design Criteria Pp 186 (Cambridge, Mass and London The MIT Press, 1976) $14 95

PUGH, III, Alexander L Dynamo User s Manual Pp x + 131 ISBN-0-262-66029-6 (Cambridge, Mass and London The MIT Press, 1976) $6 95

ROSENFELD, Azriel, and KAK, Avinash C Digital Picture Processing (Computer Science and Applied Mathematics a Series of Monographs and Textbooks) Pp xin + 457 ISBN-0-12-597360-8 (New York and London, Academic Press, a Subsidiary of Harcourt Brace Jovanovich, Publishers, 1976) $28 50 £17 40

RUBINOFF, Morris, and YOVITS, Marshall C, (edited by) Advances in Computers, Vol 15 Pp xin + 301 ISBN-0-12-012115-8 (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich, Publishers, 1976) $28 50 £17 40

SIMON, J C (edited by) Computer Oriented Learning Processes (NATO Advanced Study Institutes Series Series E Applied Science, No 14) Pp 595 ISBN-90-286-0475-8 (Leyden Noordhoff, 1976) Dfl 110. S41 50

VENTRY, Lord, and KOLESNIK Eugene M Jane's Pocket Book of Airship Development Pp 244 ISBN-0356-04656-7 (London Macdonald and Jane s Publishers, Ltd, 1976) Cloth £3 75 Limp PVC £2 75

Earth sciences

AOKI, Hitoshi, and IIZUKA, Susumu (edited by) Volcanoes and Tectonosphere Pp xxi + 370 (Tokyo Tokai University Press 1976) $40

ATTEWELL, P B, and FARMER, I M Principles of Engineering Geology Pp xxx + 1045 ISBN-0-470-03641-9 (London Chapman and Hall, New York Halsted Press a Division of John Wiley and Sons, Inc, 1976) £25 net

BENOIT, Claude Les Ressources des Oceans Alimentaires, Energetiques, Minerales Pp 98 (Paris Eyrolles, 1976) np

BUCHWALD, V F Handbook of Iron Meteorites Their History, Distribution, Composition and Structure Vol 1 Iron Meteorites in General—Tables, Appendices, References Pp xin + 1-254 + 8 Vol 2 Iron Meteorites–Abakan–Mejillones Pp 245-820 + 8 Vol 3 Iron Meteorites–Mercedritas–Zerhamra, Supplement Pp 821-1418 + 8 ISBN-0-520-02934-8 (Berkeley, Los Angeles and London University of California Press, 1975. Published for The Center for Meteorite Studies, Arizona State University) £98

DREGNE H E Soils of Arid Regions Development in Soil Science 6 Pp xin + 237 ISBN-0-444-41439-8 (Amsterdam, Oxford and New York Elsevier Scientific Publishing Company, 1976) Dfl 75 00, $28 95

ESCHER, Arthur, and WATT, W Stuart (edited by) Geology of Greenland Pp 603 ISBN-87-980404-0-5 (Copenhagen Grønlands Geologiske Undersøgelse, The Geological Survey of Greenland, Østervoldgade 10, 1976) DKr 195 including surface postage

HAY, Richard L Geology of the Olduvai Gorge a Study of Sedimentation in a Semiarid Basin Pp xvi + 203 ISBN-0-520-02963-1 (Berkeley, Los Angeles and London University of California Press, 1976) £15 75

JERLOV, N G Marine Optics (Elsevier Oceanography Series, 14) Second revised and enlarged edition of Optical Oceanography Pp xin + 231 ISBN-0-444-41490-8 (Amsterdam, Oxford and New York Elsevier Scientific Publishing Company, 1976) Dfl 77, $29 75

MEANS W D Stress and Strain Basic Concepts of Continuum Mechanics for Geologists Pp xi + 339 ISBN-3-540-07556-9 (Berlin and New York Springer-Verlag, 1976) DM 36 20, $14 80

POUQUET, Jean Earth Sciences in the Age of the Satellite Pp ix + 169 ISBN-90-277-0437-6 (Dordrecht and Boston, USA D Reidel Publishing Company, 1974) Dfl 58, $22 50

RIXON, A E Fossil Animal Remains Their Preparation and Conservation Pp vi + 304 ISBN-0-485-12028-3 (London The Athlone Press, University of London, 1976 Distributed by Tiptree Book Services, Ltd, Tiptree, Essex, Humanities Press, Inc, New Jersey) £5 25

STRONG, D F (edited by) Metallogeny and Plate Tectonics (Geological Association of Canada Special Paper No 14) Pp vn+660 (Waterloo, Ontario Geological Association of Canada, Department of Earth Sciences, University of Waterloo, 1976) np

WALTER, M R (edited by) Stromatolites (Developments in Sedimentology, 20) Pp xi+790 ISBN-0-444-41376-6 (Amsterdam, Oxford and New York Elsevier Scientific Publishing Company, 1976) Dfl 259, $99 95

Biological sciences

ANDREWS, John Birds and Their World Pp 192 ISBN-0-600-34883-0 (London and New York Hamlyn Publishing Group, Ltd, 1976) £3 95

BAKER, Joseph T, and MURPHY, Vrem Handbook of Marine Science Vol 1 Compounds from Marine Organisms Pp 226 ISBN-0-87819-391-X (Cleveland Ohio CRC Press, Inc, 1976) £25 60

BALDREY, P The Battle against Bacteria A Fresh Look Pp 179 ISBN-0-521-21268-5 (Cambridge, London, New York and Melbourne Cambridge University Press, 1976) £4 50

BARONDES, Samuel (edited by) Neuronal Recognition Pp xvi+367 ISBN-0-412-21430-X (London Chapman and Hall, Ltd, 1976) £20

BERNHEIMER, Alan W Mechanisms in Bacterial Toxicology (Developments in Medical Microbiology and Infectious Diseases) Pp xi+263 ISBN-0-471-07105-6 (New York and London John Wiley and Sons, 1976) £13 80 $23 40,

BLASIUS, W Problems of Life Research Physiological Analyses and Phenomenological Interpretations Pp xiv+197 ISBN-3-540-07731-6 (Berlin and New York Springer-Verlag, 1976) DM 19 80, $8 20

BLESCHER, Melvin (edited by) Methods in Receptor Research, Part 1 (Methods in Molecular Biology) Pp xv+383 ISBN-0-8247-6414-5 (New York and Basel Marcel Dekker, Inc, 1976) SFr 120

BOHM, Anton Morphologische Studien Im Bereiche der Pyrrhophyta Das Problem Form und Selektion (Bibliotheca Phycologica, Band 22) Pp 119 (32 tafeln) ISBN-3-7682-1069-3 (Vaduz J Cramer, 1976) DM 50

BRADLEY, P B, and DHAWAN, B N (edited by) Drugs and Central Synaptic Transmission Pp 391 ISBN-333-12735-8 (London and Basingstoke, The Macmillan Press, Ltd, 1976) £15 net

BRAUER, K and SCHOBER, W Catalogue of Mammalian Brains (Jena, Gustav Fischer Verlag, VEB, Villengang 2, Postfach 176, DDR-69 Jena, 1976) Np

Applied biological sciences

AMAVIS, R, and SMEETS, J (edited by) Principles and Methods for Determining Ecological Criteria on Hydrobiocenoses (Proceedings of the European Colloquium, Luxembourg, November, 1975) Pp 531 ISBN-0-08-021233-6 (Oxford and New York Pergamon Press, 1976 Published for the Commission of the European Communities) Hardcover £17 50, $35

ARNOLD, M H (edited by) Agricultural Research for Development. The Namulonge Contribution (Cotton Research Corporation) Pp x+353 ISBN-0-521-21051-8 (Cambridge, London and New York Cambridge University Press, 1976) £13

BERGSMA, Daniel, BRON, Anthony J, and COTLIER, Edward (edited by) The Eye and Inborn Errors of Metabolism (Symposium held at the Radcliffe Infirmary, Oxford, April 1975 Sponsored by the National Foundation—March of Dimes) Pp xxi+605 ISBN-0-8451-1003-9 (New York Alan R Liss, Inc, 1976) $60 00

VAN BEVER, Willem, and LAL, Harbans (edited by) Synthetic Antidiarrheal Drugs Synthesis—Preclinical and Clinical Pharmacology (Modern Pharmacology-Toxicology a Series of Monographs and Textbooks) Pp xiv+284 ISBN-0-8247-6370-X (New York and Basel Marcel Dekker, Inc, 1976) SFr 100

BOSMA, James F, and SHOWACRE, Jane (edited by) Symposium on Development of Upper Respiratory Anatomy and Function Implications for Sudden Infant Death Syndrome Pp x+279 (Washington, DC US Government Printing Office, 1976) $7 10

CALVIN, Melvin, and GAZENKO, O G (general editors) Foundations of Space Biology and Medicine (Joint USA/USSR Publication Vol 1 Space as a Habit Pp xvm+442 $11 75 Vol 2, Book 1 Ecological and Physiological Bases of Space Biology and Medicine Pp x+405 $9 40 Vol 2, Book 2 Ecological and Physiological Bases of Space Biology and Medicine Pp x+406-756 $8 80 Vol 3 Space Medicine and Biotechnology Pp x+542 $11 (Washington, DC Scientific and Technical Information Office, National Aeronautics and Space Administration, 1975 For sale by US Government Printing Office, Washington)

CLARKE, Frank H (editor-in-chief) Annual Reports in Medicinal Chemistry, Vol 11 Pp ix+330 ISBN-0-12-040511-3 (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich, Publishers, 1976) $16 50, £10 05

COLLEE, J G Applied Medical Microbiology (Basic Microbiology, Vol 3) Pp 121 ISBN-0-632-00718-3 (Oxford and London Blackwell Scientific Publications, 1976 Distributed in the USA by Halsted Press, a Division of John Wiley and Sons, Inc, New York) £2 80

COPELAND, L O Principles of Seed Science and Technology Pp xiv+369 ISBN-0-8087-0361-7, (Minneapolis, Minnesota Burgess Publishing Company, 1976) $14 95.

DEBRY, G, and BLEYER, R (edited by) Alimentation et Travail (Deuxieme Symposium International 15-16-17 Mai 1974, A Vittel (France) Organise par La Societe de Nutrition et de Dietetique de Langue Francaise) Pp 550 (Nancy Vagnor, 1976) np

FAJANS, Stefan S (edited by) Diabetes Mellitus Pp xm+361 (Bethesda, Maryland Department of Health Education and Welfare, National Institutes of Health, 1976 For sale by U S Government Printing Office, Washington, D C) np

HAYAT, M A (edited by) Principles and Techniques of Scanning Electron Microscopy Biological Applications Vol 2 Pp xn+171 ISBN-0-442-25678-7 (New York, London, Toronto and Melbourne Van Nostrand Reinhold Company, 1974) £13 70

HUDSON, L, and HAY, F C Practical Immunology Pp xvi+298 ISBN-0-632-00211-5 (Oxford, London, Edinburgh and Melbourne Blackwell Scientific Publications, 1976) £6 50

KEEN, H, JARRETT, J, and LEVY, A M Triumphs of Medicine Pp 193 ISBN-0-236-30925-0 (London and New Hampshire Paul Elek, 1976) £12 50

RIGHETTI, P G, and DRYSDALE, J W Isoelectric Focusing (Laboratory Techniques in Biochemistry and Molecular Biology, Volume 5 Part 11 Pp 337-590 ISBN-0-7204-4219-2 (Amsterdam and Oxford North-Holland Publishing Company, New York American Elsevier Publishing Co, Inc, 1976) $13 95, Dfl 36 00

ROSA, N R, and FRIEDMAN, H (edited by) Manual of Clinical Immunology Pp xiv+932 ISBN-0-914826-09-3 (Washington American Society for Microbiology, 1976) Hardback $20, Flexibleback $16

VERRIEST, Guy (edited by) Colour Vision Deficiencies III (Proceedings of the Third Symposium of the International Society on Colour Vision Deficiencies, Amsterdam/Leiden June 25-27, 1975) (Modern Problems in Ophthalmology, Vol 17) Pp vn+380 ISBN-3-8055-2297-5 (Basel, Paris and New York S Karger, 1976) Sf/DM 178

WIGGINS, G S, and WILSON, A A Colour Atlas of Meat and Poultry Inspection ISBN-0-7234-0701-8 (London Wolfe Medical Publications Limited, 10 Earlham Street, London WC2, 1976) £17 50

RESEARCH ON THE FETUS (The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research) Report and Recommendations Pp 88 Appendix (Bethesda, Md US Department of Health, Education and Welfare, 1976)

Anthropology

BOLTON, M Etheopean Wildlands Pp 221 ISBN-0-00-262205-X (London Collins and Harvill Press, 1976) £5 50

WENDORF, F, and SCHILD, R Prehistory of the Nile Valley Pp xxv+404 ISBN-0-12-743950-1 (New York, Academic Press, a subsidiary of Harcourt Brace Jovanovich, Publishers, 1976) $25 00, £13 75

Psychology

BRIERLY, J The Growing Brain Childhood's Crucial Years Pp 138 ISBN-0-85633-099X) (Windsor NFER Publishing Company, 1976) £3 55

EYSENCK, H J, and WILSON, G D (edited by) The Experimental Study of Freudian Theories Pp xv+405 ISBN-0-416-78010-5 (London Methuen and Company Ltd, 1976) £3 90

GILLAN, P & R Sex Therapy Today Pp 234 ISBN-0-7291-0045-6 (London Open Books, 1976) £5 95

MATTILA, M (edited by) Alcohol, Drugs and Driving Pp vn+101 ISBN-3-8055-2341 (Basel, Munich, Paris, London, New York and Sydney S Karger, 1976) DM 49, $19 00

MORRIS, J D, ROLL, W G, and MORRIS, R L (edited by) Research in Parapsychology 1975 (Abstracts and Papers from the Eighteenth Annual Convention of the Parapsychological Association, 1975) Pp vn+277 ISBN-0-8108-0895-1 (Metuchen, N J The Scarecrow Press, Inc Folkestone Bailey Brothers and Swinfen Ltd, 1976) £7 20

SCHMEIDLER, Gertrude R (edited by) Parapsychology Its Relation to Physics, Biology, Psychology and Psychiatry Pp xm+278 ISBN-0-8108-0909-5 (Metuchen, N J The Scarecrow Press, Inc, Folkestone Bailey Brothers and Swinfen Ltd, 1976) £3 90

TEMERLIN, M K Lucy Growing up Human A Chimpanzee Daughter in a Psychotherapy Family Pp xxn+216 ISBN-0-285-62238-2 (London Souvenir Press, 1976) £3 50

THIESSEN, D D The Evolution and Chemistry of Aggression Pp xvn+211 ISBN-0-03455-9 (Springfield, Illinois Charles C Thomas, 1976) $18 50

VOLOSINOV, V N Freudianism A Marxist Critique Pp xvn+153 ISBN-0-12-723259 (New York, San Francisco and London Academic Press, a subsidiary of Harcourt Brace Jovanovich, Publishers, 1976) $10 95 £6 70

WAUQUIER, A, and ROLLS, E T (edited by) Brain-Stimulation Reward Pp xm+621 ISBN-0-7204-0587-4 (Amsterdam and Oxford North-Holland Publishing Company, New York American Elsevier Publishing Company, Incorporated, 1976) $56 95, Dfl 148 00

History and philosophy of science

HALL, D H History of the Earth Sciences during the Scientific and Industrial Revolutions with special emphasis on the physical geosciences Pp xi+297 ISBN-0-444-41440-1 (Amsterdam, Oxford and New York Elsevier Scientific Publishing Company, 1976) Dfl 50 00, $19 25

LINDBERG, D C Theories of Vision from Al-Kindi to Kepler Pp xin+324 ISBN-0-226-48234-0 (Chicago and London University of Chicago Press, 1976) £13 60

MORAZE, C, (edited by) History of Mankind Cultural and Scientific Development Volume Five, The Nineteenth Century I & II, The Scientific Revolution, Industrial Revolution and Technical Developments Pp xxix+429 ISBN-3-04-900026-8 III, Social, Cultural and Religious Aspects Pp xi+433-1009 ISBN-0-04-900027-6 IV, European Empires, Technical and Scientific Progress, Culture Conflicts Pp xi+1013+1394 ISBN-0-04-900028-4 (London George Allen and Unwin Limited, 1976) £15 00 per volume

SPICKER, S F, and ENGELHARDT, H T Jr Philosophy and Medicine Volume Philosophical Dimensions of the Neuro-Medical Sciences Pp vi+274 ISBN-90-277-0672-3 (Dordrecht, Holland and Boston D Reidel Publishing Company, 1976) np

General

ANANICHEV, K Environment International Aspects Translated from the Russian by Yuri Shvirkov Pp 207 (Moscow Progress Publishers, 1976 Distributed in Great Britain by Central Books, Ltd, 37 Grays Inn Road, WC1) 75p

ARMAMENTS AND DISARMAMENT IN THE NUCLEAR AGE a Handbook Pp 308 ISBN-0-391-00652-5 (Stockholm Stockholm International Peace Research Institute, 1976 Published in collaboration with Almqvist and Wiksell International. Stockholm and Humanities Press, Inc, Atlantic Highlands, NJ) $4 Kr 82

BAXTER, John, and ATKINS, Thomas The Fire Came By The Riddle of the Great Siberian Explosion Pp 165+39 plates ISBN-0-354-04012-X (London Macdonald and Jane s, 1976) £3 95

BOULLOCHE, Andre, RICHTER, Klaus, and WARREN, Kenneth (compiled and presented by) The Sciences and Democratic Government Highlights of the Fourth Parliamentary Scientific Conference, Florence, 1975 (Parliamentary Assembly of the Council of Europe Committee Science and Technology) Pp x+149 ISBN-0-333-19871-9 (London and Basingstoke The Macmillan Press, Ltd, 1976) £6 95

BRAIN, R The Tribal Impulse A New Look at Traditional Communities Pp 288 ISBN-0-354-04018-9 (London Macdonalds and Jane s, 1976) £6 95

BRAND, Myles, and WALTON, Douglas (edited by) Action Theory (Proceedings of the Winnipeg Conference on Human Action, held at Winnipeg, Manitoba, Canada 9-11 May 1975) (Synthese Library, Vol 97) Pp vi+345 ISBN-90-277-0671-9 (Dordrecht and Boston USA D Reidel Publishing Company, 1976) Dfl 105, $39 50

CHRISTIE, Clifford And Then They Fly Away Pp 159+17 photographs ISBN-0-09-461110-6 (London Constable and Co, Ltd 1976) £3 95 net

DAVIES, Duncan, BANFIELD, Tom, and SHEAHAN, Ray The Humane Technologist (Science and Engineering Policy Series) Pp vn+180 ISBN-0-19-858325-7 (London Oxford University Press, 1976) £4 75 net

DINELEY, David, HAWKES, Donald, HANCOCK, Paul, and WILLIAMS, Brian Earth Resources, a Dictionary of Terms and Concepts Pp 206 ISBN-0-09-913690-2 (London Arrow Books, Ltd, 1976) £1 25

DOUGLAS, Alfred Extra-Sensory Powers a Century of Psychical Research Pp 392+14 photographs ISBN-0-575-02137-3 (London Victor Gollancz, Ltd, 1976) £6 95 net

DRABBLE, Phil Country Seasons Pp 213 ISBN-0-7181-1548-1 (London Michael Joseph 1976) £4 25 net

FLADER, Susan L Thinking Like a Mountain Aldo Leopold and the Evolution of an Ecological Attitude Toward Deer, Wolves, and Forests Pp xvi+284 ISBN-0-8262-0167-9 (Columbia and London University of Missouri Press, 1974) £8 75

GILLIE, Oliver, Who Do You Think You Are? Man or Superman The Genetic Controversy Pp 255 ISBN-0-246-10836-3 (London Hart-Davis, MacGibbon, Ltd, 1976) £4 95 net

HOLISTER Geoffrey, and PORTEOUS, Andrew The Environment a Dictionary of the World Around Us Pp 280 ISBN-0-09-913470-5 (London Arrow Books, Ltd, 1976) £1 25

HUMBER, James M, and ALMEDER, Robert F (edited by) Biomedical Ethics and the Law Pp xi+541 ISBN-0-306-30902-5 (New York and London Plenum Press 1976) $23 40

JANTSCH, Erich, and WADDINGTON, Conrad H (edited by) Evolution and Consciousness Human Systems in Transition Pp xii+259 ISBN-0-201-03439-5 (Reading Mass and London Addison-Wesley Publishing Company, 1976) Cloth $19 50, Paper $9 50

JEVONS, F R Knowledge and Power (The 1975 Lectures of The Australian National University Advisory Committee on Science and Technology Policy Research) Pp 111 ISBN-0-7081-0333-2 (Canberra Australian National University Press, 1976) A$4 95

KENNET Wayland (edited by) The Futures of Europe (Based on a Report to the Commission of the European Communities) Pp xii+242 ISBN-0-521-21326-6 (Cambridge London and New York Cambridge University Press, 1976) £6 50

KERSHAW David, and FAIR Jerilyn The New Jersey Income-Maintenance Experiment Vol 1 Operations, Surveys, and Administration (Institute for Research on Poverty Monograph Series) Pp xvn—232 ISBN-0-12-405001-8 (New York and London Academic Press a Subsidiary of Harcourt Brace Jovanovich, Publishers 1976) $15 £9 15

LIN, Steven A Y (edited by) Theory and Measurement of Economic Externalities (Economic Theory and Mathematical Economics) Pp xiv+265 ISBN-0-12-450450-7 (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich, Publishers 1976) $19 50 £11 90

McHALE, John The Changing Information Environment (Environmental Studies) Pp 117 ISBN-0-236-40052-5 (London Paul Elek 1976) £2 95 net

MEASURES, David G Bright Wings of Summer Watching Butterflies Pp 160, ISBN-0-304-29764-X (London Cassell and Co Ltd, 1976) £5 net

SCIENCE YEAR The World Book Science Annual, 1977 (A Review of Science and Technology During the 1976 School Year) Pp 431 ISBN-0-7166-0577-5 (Chicago and London Field Enterprises Educational Corporation 1977) np

SEYMOUR, John The Complete Book of Self-Sufficiency Pp 256 ISBN-0-571-11095 (London Faber and Faber, Ltd 1976) £5 50 net

SHRYOCK, Henry S, SIEGEL, Jacob, S and Associates Condensed Edition by Edward Stockwell The Methods and Materials of Demography (Studies in Population) Pp ix+577 ISBN 0-12-641150-6 (New York and London Academic Press, a Subsidiary of Harcourt Brace Jovanovich, Publishers, 1976) $16 50, £9 55

SIEGHART, Paul Privacy and Computers Pp vn+228 ISBN-901539-56-2 (London Latimer New Dimensions Limited, 1976) £6 50

SUTHERLAND Stuart Breakdown a Personal Crisis and a Medical Dilemma Pp x+274 ISBN-0-297-77217-1 (London Weidenfeld and Nicolson, 1976) £5 25 net

TETT Norman, and CHADWICK, John (edited by) Professional Organizations in the Commonwealth Revised edition Pp xi+584 ISBN-0-09-127240-8 (London Hutchinson and Co (Publishers), Ltd, 1976 Published for The Commonwealth Foundation) £10